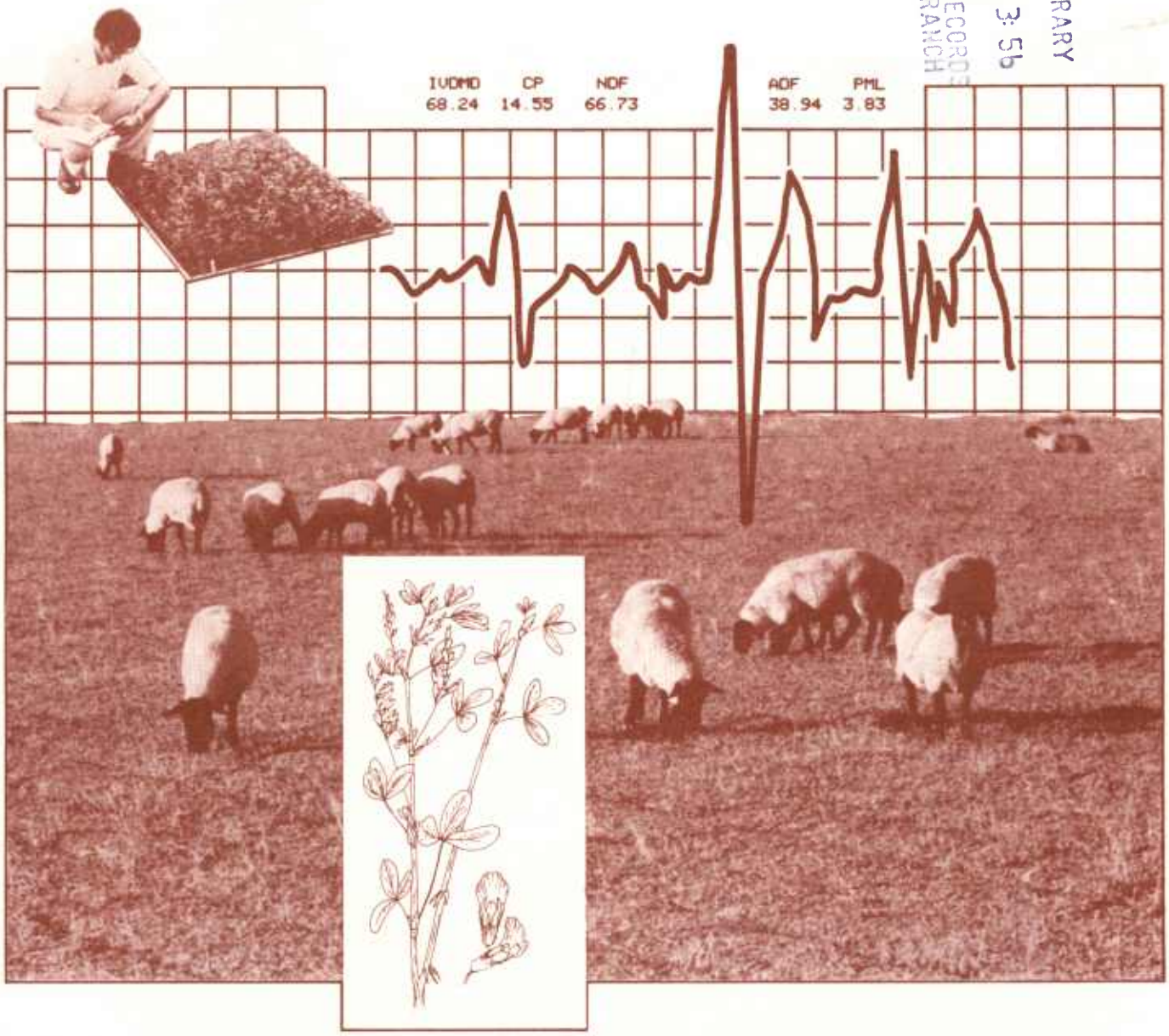
United States
Department of
Agriculture

Agricultural
Research
Service

Agriculture
Handbook
No. 643

Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality



Abstract

Marten, G. C., J. S. Shenk, and F. E. Barton II, editors. 1989. Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality. U.S. Department of Agriculture, Agriculture Handbook No. 643 (revised with supplements), 110 p.

Near infrared reflectance spectroscopy (NIRS) is a nonconsumptive, instrumental method for fast, accurate, and precise evaluation of the chemical composition and associated feeding value attributes of forages and other feedstuffs provided the proper procedures are followed. Each of the major organic feed components has absorption characteristics (due to vibrations arising from the stretching and bending of H bonds associated with C, O, and N) in the near infrared region that are specific to the component. These absorption characteristics primarily determine diffuse reflectance, which gives us the means for assessing composition. This handbook is intended to provide an overview of the history of NIRS use for forage quality analysis, a summary of available equipment and computer software, recommended conventional and NIRS procedures for analysis of forage quality, an explanation of how and why NIRS functions, recommended NIRS instrument calibration (including equation development, validation, monitoring and transfer), progress in NIRS technology transfer, and consideration of future applications and certification procedures for NIRS technology. Readers may find that specific chapters meet their informational needs, or they may wish to read all chapters, appendixes, and supplements sequentially for a complete interpretation of the subject.

Keywords: Feedstuff evaluation, forage evaluation, forage feeding value, forage quality, forage sampling, near infrared reflectance, NIR spectroscopy, NIRS analysis, NIRS calibration, NIRS equations, NIRS equipment, NIRS software, NIRS spectra, NIRS technology transfer.

United States
Department of
Agriculture

Agricultural
Research
Service

Agriculture
Handbook
No. 643

Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality

Editors:
G.C. Marten
J.S. Shenk
F.E. Barton II

Revised August 1989 (with supplements)

Contributors

S. M. Abrams
Research Animal Scientist
USDA-ARS
U.S. Regional Pasture Research Laboratory
U.S. Dairy-Forage Research Center-Pennsylvania Cluster
University Park, Pennsylvania 16802

F. E. Barton II
Research Chemist
USDA-ARS
Richard B. Russell Agricultural Research Center
P.O. Box 5677
Athens, Georgia 30613

T. H. Blosser
Research Animal Scientist
USDA-ARS
Ruminant Nutrition Laboratory
BARC-East, Building 200
Beltsville, Maryland 20705

D. H. Clark
Research Animal Scientist
USDA-ARS
Skaggs Research Building, UMC-94
Utah State University
Logan, Utah 84322

S. W. Coleman
Research Animal Nutritionist
USDA-ARS
Livestock and Forage Research Laboratory
P.O. Box 1199
El Reno, Oklahoma 73036

J. G. Linn
Extension Animal Scientist-Dairy Nutrition
Animal Science Department
University of Minnesota
St. Paul, Minnesota 55108

G. C. Marten
Research Agronomist
USDA-ARS
Agronomy and Plant Genetics Department
University of Minnesota
St. Paul, Minnesota 55108

N. P. Martin
Extension Agronomist-Forages
Agronomy and Plant Genetics Department
University of Minnesota
St. Paul, Minnesota 55108

D. R. Mertens
Research Dairy Scientist
USDA-ARS
U.S. Dairy Forage Research Center
University of Wisconsin
Madison, Wisconsin 53706

K. H. Norris
Research Agricultural Engineer
USDA-ARS
Instrumentation Laboratory
BARC-West, Building 002
Beltsville, Maryland 20705

J. S. Shenk
Professor
Agronomy Department
The Pennsylvania State University
University Park, Pennsylvania 16802

W. C. Templeton, Jr.
Research Agronomist (Retired)
USDA-ARS
U.S. Regional Pasture Research Laboratory
University Park, Pennsylvania 16802

M. O. Westerhaus
NIR Program Analyst
Agronomy Department
The Pennsylvania State University
University Park, Pennsylvania 16802

W. R. Windham
Research Animal Physiologist
USDA-ARS
Richard B. Russell Agricultural Research Center
P.O. Box 5677
Athens, Georgia 30613

Contents

Definition of NIRS Analysis

K. H. Norris

NIRS method of analysis is an instrumental method for rapidly and reproducibly measuring the chemical composition of samples with little or no sample preparation. It is based on the fact that each of the major chemical components of a sample has near infrared absorption properties which can be used to differentiate one component from the others. The summation of these absorption properties, combined with the radiation-scattering properties of the sample, determines the diffuse reflectance of a sample. Therefore, the near infrared diffuse reflectance signal contains information about the composition of the sample. The compositional information can be extracted by proper treatment of the reflectance data. The simplest procedure is to measure the reflectance at two wavelengths, with one wavelength chosen to be at a maximum absorption point and the other chosen to be a minimum absorption point, for the constituent to be analyzed. The ratio of these two reflectance values measured on different samples can be correlated to the concentration of that specific constituent in those samples. By performing this correlation, an equation can be developed to predict the concentration of the constituent in samples from their reflectance measurements. Such equations can be developed for each of the constituents in the samples by using reflectance data at selected wavelengths. There are many possible methods of treating the reflectance data to obtain the best possible prediction of the sample composition. These will be discussed in a later chapter.

The NIRS method of analysis has four main advantages: speed, simplicity of sample preparation, multiplicity of analyses with one operation, and nonconsumption of the sample (so that it can be analyzed again by the same or another procedure). The NIRS measurement can be made in less than 1 second, although a more typical time is from 30 seconds to 3 minutes. Usually, the sample preparation consists only of grinding to a mean particle size of from 100 to 500 μm to ensure sample homogeneity. The ground sample is placed in the sample cup and is ready for measurement. By making measurements at many wavelengths, many constituents are measured at the same time.

The chief disadvantages of the NIRS method are instrumentation requirements, dependence on calibration procedures, complexity in the choice of data treatment, and lack of sensitivity for minor constituents. The NIRS method requires high-precision spectroscopic instrumentation because small changes in reflectance at specific wavelengths must be measured. Calibration is required for each constituent, and, in general, a calibration is valid only for the same type of samples. Different instrument manufacturers use different data treatment procedures, and various proponents of the NIRS method do not agree on the optimum data treatment.

The technology of the NIRS method is still in the developmental stage. We can anticipate that changes will occur; but at the present state of development, it is being widely used in the grain and feed industry. For forages, the NIRS method of analysis makes it possible to take a sample from a truckload of hay and provide, in less than 3 minutes, an analysis for crude protein, acid detergent fiber, neutral detergent fiber, dry matter, lignin, and in vitro dry matter digestibility.

History of NIRS Analysis of Agricultural Products

D. H. Clark

Research on NIRS has been quite extensive over the last 10 years. Basically, research has been with either filter-type instruments or scanning monochromators (SM). Results have varied, depending upon type of instrument, analysis performed, and statistical treatment of data generated. Statistical interpretations associated with NIRS analyses are explained by Shenk et al. (*1978*) and elsewhere in this handbook.

Filter Instruments

Cereals and Oil-Bearing Seeds

Many researchers have used NIRS analyses to determine the concentrations of moisture, crude protein (CP), and oil in cereal grains and oil-bearing seeds. Williams (*1975*) reported that standard errors of difference in percentage of moisture were from 0.12 for barley to 0.30 for rapeseed. He also examined some of the sources of error (sampling, instrument, sample packing, and so forth) associated with NIRS. The standard errors of reproducibility for CP and moisture were large when he tested freshly ground samples. Standard errors of duplicate moisture or oil values (samples were scanned, rotated, and rescanned) were comparable to those associated with the chemical analysis. Crude protein calibration errors were reduced by increasing the range for calibration purposes from 12–14 to 10–18 percent protein. He also noted that standard errors of deviation between duplicate readings were reduced from 0.30 to 0.15 percent by mixing the samples 15 times.

Rubenthaler and Bruinsma (*1978*) analyzed lysine in wheat samples and reported that the standard error of analysis (SEA) was 0.08 mg/100 g CP. In addition, Gill et al. (*1979*) analyzed the concentrations of nitrogen (N) and lysine in barley; they found the results sufficiently accurate for preliminary selection purposes in breeding projects.

Watson et al. (*1976*) compared three NIRS grain analyzers (one with a fixed filter and two with a scanning filter) and three grinders (cyclone mill, coffee mill, and a modified pulverizer) for analyzing protein concentrations in wheat. Differences in grinders affected the results more than did differences in NIRS instruments. Instrument differences were evident when Hunt et al. (*1977*) analyzed oil in soybeans. Two scanning filter instruments were compared for analyzing moisture in whole kernels of corn and sorghum (Stermer et al. *1977*). The standard errors of estimate (SEE) were 0.8 and 3.4 percent, respectively, for wet samples of corn and sorghum. Instrument differences were evident and were attributed to differences in the number of readings averaged for each instrument. In preparing cornmeal from different kernel types, Hymowitz et al. (*1974*) found that increasing the grinding time decreased the CP readings and increased the oil readings.

Starr et al. (*1981*) successfully analyzed N in such diverse commodities as spring field beans, winter triticale, spring and winter wheats, rapeseed, kale leaves and stems, and spring and winter barleys with a fixed-filter instrument. They were also able to calibrate their instrument to indicate the grinding resistance of wheat kernels and extent of bran removal—two indexes of baking quality.

Forages and Animal Response

Use of NIRS filter instruments for analyses of forage quality and animal response has been reported by several investigators. Winch and Major (*1981*) reported low standard errors of calibration (SEC) and SEA when analyzing N in grasses, legumes, and legume-grass mixtures. However, analyses for in vitro and in vivo dry matter digestibility resulted in large errors. Starr et al. (*1981*) noted that the small range in variation and large SEAs prevented NIRS instruments equipped with 19 filters from being useful for analyzing N in wheat straw.

Counts and Radloff (*1979*) used an instrument containing six filters to analyze CP and in vitro dry matter digestibility (IVDMD) in alfalfa, grasses, and alfalfa-grass mixtures. They reported SEAs comparable to those of conventional laboratory analyses. They also reported an r^2 value of 0.83 between NIRS-estimated dry matter digestibility and results from metabolism trials. Barton and Burdick (*1983*), using a scanning-filter instrument, obtained SEC and SEA values of 1.78 and 2.54 percent, respectively, for digestible dry matter in bermudagrass hays. Fales and Cummins (*1982*) examined the effects of storing forage-type sorghum samples under different humidities on NIRS analyses of acid detergent fiber (ADF) content; they noted that the SEEs for the high-humidity samples were as much as two times the SEEs for the low-humidity samples.

To reduce the bias in CP, in vivo DMD, and voluntary intake analyses made with a fixed-filter NIRS instrument, Minson et al. (*1983*) reported that differences in grass species, plant parts, and form of processing must be taken into account. Branine et al. (*1983*) found inflated standard errors for CP and IVDMD when shrubs and forbs or a combination of shrubs, forbs, and grasses

(compared to grasses alone) were analyzed by a fixed-filter instrument. Barton and Burdick (*1979*) reported that separate equations should be used with a tilting-filter instrument to analyze warm and cool season grasses for CP, ADF, neutral detergent fiber (NDF), lignin, and IVDMD. Burdick et al. (*1981*) found that NIRS equations for the analysis of CP in bermudagrass (freeze dried or sun cured) did not apply well to drum-dehydrated bermudagrass pellets.

Meat and Dairy Products

Kruggel et al. (*1981*) analyzed meat products with an NIRS instrument equipped with six filters. They found that a high fat content and changes in meat temperature caused the biggest variation in analyses of beef, lamb, and pork. They were successful in analyzing fat and moisture in beef and lamb samples that had a low fat content.

Giangiacomo et al. (*1979*) analyzed freeze-dried and powdered blue cheese. They reported correlations of 0.93, 0.91, 0.90, 0.81, and 0.70 between NIRS and laboratory analyses of protein, ash, free tyrosine, pH, and fat, respectively. However, they failed to adquately measure mold counts, standard plate counts, and gram-negative bacteria by NIRS. White et al. (*1978*) successfully analyzed fat and total solids in raw milk, pasteurized whole milk, lowfat milk, skim milk, ice milk mix, ice cream mix, and fluid chocolate products with a fixed-filter intrument. Lactose, fat, and moisture in powdered milk have also been adquately analyzed by using NIRS (Rosenthal *1977*).

Other Agricultural Products

Rosenthal (*1977*) reported that analyses made on other agricultural products by USDA with NIRS filter instruments include mold damage in corn, maturity of peaches and apples, blood in eggs, smut content of wheat, color of tomatoes, and degree of milling of rice.

Scanning Monochromators

Scanning monochromators (SM) are important tools for NIRS research in that the whole spectrum within the near-infrared region is utilized. This allows for the analysis of chemical constituents that may be less adequately assayed with filter instruments.

Cereals and Oil-Bearing Seeds

Ben-Gera and Norris (*1968*) found no difference in the moisture content of soybean flour assayed by SM-NIRS and by an oven-drying method. Shenk and Barnes (*1977*) noted that similar wavelengths were used for analyzing CP in grains and forages. However, different wavelengths were needed for analyzing IVDMD in these feeds. Calibration for CP in grains was more accurate within species than across species (Shenk and Barnes *1977*). Standard errors (percent of the total sample) during SM-NIRS analyses of 17 amino acids in wheat ranged from 0.015 to 0.183 for histidine and glutamic acid, respectively (Rosenthal *1977*).

Forages

Norris et al. (*1976*) reported SEAs of 0.95, 3.1, 2.5, 2.1, and 3.5 percent, respectively, for CP, NDF, ADF, lignin, and IVDMD in a diverse mixture of forages. In that publication, the first on using NIRS for forage quality analysis, they also reported that standard errors were reduced when they used fine particles rather than coarse ones. Shenk et al. (*1981*) also used an SM-NIRS instrument to analyze forages representing a diversity of species, mixtures, maturities, and locations; they reported standard errors of 1.13, 1.27, 0.16, 0.04, and 0.37 percent for lignin, cellulose, calcium (Ca), phosphorus (P), and potassium (K), respectively. In addition, Shenk et al. (*1981*) found that errors for CP, IVDMD, and fiber fractions were similar in size to those reported by Norris et al. (*1976*). The authors noted that log 1/R (R = reflectance) was the best mathematical treatment of the spectra for CP and IVDMD and that the second derivative of the spectra was the best for analyzing the fibrous components and minerals. Thirty forage samples were scanned with the SM at six USDA locations (Templeton et al. *1983*). Spectra from the instruments (same make and model) were identical except in the moisture region (1940 nm).

Marten et al. (*1983*) reported acceptable standard errors when they analyzed quality components in spring oat, barley, wheat, and triticale forages with an SM-NIRS instrument. Standard errors of analyses for total N and soluble N in sodium chloride solution were both 0.04 percent for crested wheatgrass samples (Park et al. *1983*a); SEAs for NDF, ADF, and lignin were similar to those reported by Marten et al. (*1983*). Shenk and Barnes (*1977*) noted reduced SEAs and CP and IVDMD when the whole NIR spectrum was used rather than the wavelengths recommended by commercial companies for forage analyses.

Analyses of percent legume, acid-detergent protein, boron, and Ca:P ratio in legume-grass mixtures resulted in reasonable SECs according to Shenk et al. (*1979*). Park et al. (*1983*b) noted that for carotene analyses by an SM-NIRS instrument, microwave-dried samples of baled alfalfa hay had to be calibrated separately from

samples dried by other methods. They also reported the need to split the samples into high and low (carotene) calibration sets for adequate results. Marten et al. (*1984*) stated that separate equations may give better results than one equation when analyzing different legumes and legume parts (leaves and stems), although they derived single equations for individual quality constituents that could be accurately used for total forage and plant parts of four species.

Scanning monochromator NIRS analyses of esophageal samples from forest and grassland ranges resulted in SECs (percent) and SEAs (percent) of 0.52 and 0.41 for CP and 1.44 and 1.74 for IVDMD (Holechek et al. *1982*). Ward et al. (*1982*) reported a SEC of 0.37 percent for CP in esophageal samples from arid and semiarid rangelands.

Animal Response

Lindgren (*1983*), analyzing timothy and meadow fescue, reported coefficients of variation for in vivo organic matter digestibility percent and metabolizable energy (MJ/kg dry matter) of 2.6 and 4.3 for green chop, 4.3 and 4.3 for hay, and 6.9 and 7.0 for silage samples, respectively. The need for separate equations for green chop, hay, and silage, or having an equation containing all preservation practices was noted. Shenk and Barnes (*1977*) reported lower SECs for analyses of intake and digestibility of forages with instruments capable of using the whole near infrared reflectance (NIR) spectrum than with fixed wavelength instruments.

Norris et al. (*1976*) reported SEAs of 5.1 percent and 7.9 g for dry matter digestibility and dry matter intake, respectively, when forage mixtures were analyzed by an SM-NIRS instrument. SECs of 2.3 percent and 6.1 $g \cdot day^{-1} \cdot (BW^{0.75})^{-1}$ (where BW = body weight in kilograms) for dry matter digestibility and dry matter intake were reported by Shenk et al. (*1977*). Ward et al. (*1982*) reported a SEC of 9.6 $\cdot$ $g^{-1} \cdot BW^{0.75}$ for organic matter intake when they analyzed esophageal samples. Eckman et al. (1983) used data from digestion trials with sheep fed pure and mixed forage diets to compare NIRS and conventional analyses (IVDMD, NDF, and protein) in estimating digestible energy, dry matter intake, and digestible energy intake. They noted that SECs for NIRS analyses were lower than or equal to those for conventional analyses.

Other Agricultural Products

Park et al. (*1982*) reported low standard errors for SM-NIRS analyses of CP, NDF, ash, and crude fat in such products as brussels sprouts, green beans, broccoli, cauliflower, spinach, asparagus, corn, carrots, rhubarb, onions, peas, squash, and cabbage. McClure and Weeks (*1979*) successfully analyzed total reducing sugars, alkaloids, nitrogen, nicotine, polyphenols, sterols, and water in tobacco products by SM-NIRS. The NIRS assay for yield of greasy wool using the estimated wool content of washed wool was investigated by Connell and Norris (*1980*).

This history of NIRS is not intended to be complete; it is intended to let the reader know how diverse the investigations with NIRS have been and to stimulate further investigations. Some areas needing investigation are listed under "Future Applications of NIRS."

References

Barton, F. E., II, and D. Burdick.
1983 Prediction of forage quality with NIR reflectance spectroscopy. *In* Proc. 14th Int. Grassl. Cong., Lexington, KY, p. 532-534. Westview Press, Boulder, CO.

Barton, F. E., II, and D. Burdick.
1979 Preliminary study on the analysis of forages with filter-type near infrared reflectance spectrometer. J. Agric. Food Chem. 27:1248-1252.

Ben-Gera, I., and K. H. Norris.
1968 Determination of moisture content in soybeans by direct spectrophotometry. Israel J. Agric. Res. 18:125-132.

Branine, M. E., B. F. Sowell, L. J. Krysl, and others.
1983 Estimating nutritive value of native range forage by near infrared reflectance. [abstr.] J. Anim. Sci. 57(1):279.

Burdick, D., F. E. Barton, II, and B. D. Nelson.
1981 Prediction of bermudagrass composition and digestibility with a near infrared multiple filter spectrophotometer. Agron J. 73:399-403.

Connell, J. P., and K. H. Norris.
1980 The prediction of the yield of greasy wool by near infrared reflectance spectroscopy. I. Principles and washed wool. Textile Res. J. 50:371-380.

Counts, G. E., and H. D. Radloff.
1979 The potential of infrared analysis in forage evaluation. [abstr.] J. Anim. Sci. 53(1):261.

Eckman, D. D., J. S. Shenk, P. J. Wangsness, and M. O. Westerhaus.
1983 Prediction of sheep responses by near infrared reflectance spectroscopy. J. Dairy Sci. 66:1983-1987.

Fales, S. L., and D. G. Cummins.
1982 Reducing moisture-induced error associated with measuring forage quality using near infrared reflectance. Agron J. 74:585-588.

Giangiacomo, R., D. Torreggiani, J. F. Frank, and others.
1979 Analysis of blue cheese using infrared reflectance spectroscopy. [abstr.] J. Dairy Sci. 62(1):39.

Gill, A. A., C. Starr, and D. B. Smith.
1979 Lysine and nitrogen measurement by infrared reflectance analysis as an aid to barley breeding. J. Agric. Sci. (Camb.) 93:727-733.

Holechek, J. L., J. S. Shenk, M. Vavra, and D. Arthun.
1982 Prediction of forage quality using near infrared reflectance spectroscopy on esophageal fistula samples from cattle on mountain range. J. Anim. Sci. 55:971-975.

Hunt, W. H., D. W. Fulk, B. Elder, and K. H. Norris.
1977 Collaborative study on infrared reflectance devices for determination of protein in hard red winter wheat and for protein and oil in soybeans. Cereal Foods World 22:534-536.

Hymowitz, T., J. W. Dudley, F. I. Collins, and C. M. Brown.
1974 Estimation of protein and oil concentration in corn, soybean and oat seed by near infrared light reflectance. Crop Sci. 14:713-715.

Kruggel, W. M., R. A. Field, M. J. Riley, and others.
1981 Near infrared reflectance determination of fat, protein, and moisture in fresh meat. J. Assoc. Off. Anal. Chem. 64:692-696.

Lindgren, E.
1983 The use of near infrared reflectance spectroscopy for predicting energy value of grass. Swed. J. Agric. Res. 13:229-233.

Marten, G. C., G. E. Brink, D. R. Buxton, and others.
1984 Near infrared reflectance spectroscopy analysis of forage quality in four legume species. Crop Sci. 24:1179-1182.

Marten, G. C., J. L. Halgerson, and J. H. Cherney.
1983 Quality prediction of small grain forages by near infrared reflectance spectroscopy. Crop Sci. 23:94-96.

McClure, W. F., and W. W. Weeks.
1979 Computerized NIR (near infrared reflectance) spectrophotometric measurement of the chemical composition of tobacco. (abstr.] Cereal Foods World 24:455.

Minson, D. J., K. L. Butler, N. Grummitt, and D. P. Law.
1983 Bias when predicting crude protein, dry matter digestibility and voluntary intake of tropical grasses by near infrared reflectance. Anim. Feed Sci. Tech. 9:221-237.

Norris, K. H., R. F. Barnes, J. E. Moore, and J. S. Shenk.
1976 Predicting forage quality by infrared reflectance spectroscopy. J. Anim. Sci. 43:889-897.

Park, Y. W., M. J. Anderson, A. W. Mahoney.
1982 Potential of near infrared reflectance (NIR) spectroscopy for estimating nutrient content of dehydrated vegetables. J. Food Sci. 47:1558-1561.

Park, Y. W., M. J. Anderson, K. H. Asay, and A. W. Mahoney.
1983a Predicting soluble nitrogen and fibrous fractions in crested wheatgrass with near infrared reflectance spectroscopy. J. Range Manage. 36:529-533.

Park, Y. M., M. J. Anderson, J. L. Walters, and A. W. Mahoney.
1983b Effect of processing methods and agronomic variables on carotene contents in forages and predicting carotene in alfalfa hay with near infrared reflectance spectroscopy. J. Dairy Sci. 66:235-245.

Rosenthal, R. D.
1977 An introduction to near infrared quantative analysis. 1977 Ann. Meeting Am. Assoc. Cereal Chem., p. 1-15.

Rubenthaler, G. L., and B. L. Bruinsma.
1978 Lysine estimation in cereals by near infrared reflectance. Crop Sci. 18:1039-1042.

Shenk, J. S., and R. F. Barnes.
1977 Current status of infrared reflectance. *In* 34th South Past. and Forage Crop Imp. Conf., p. 57-62.

Shenk, J. S., I. Landa, M. R. Hoover, and M. O. Westerhaus.
1981 Description and evaluation of near infrared reflectance spectro-computer for forage and grain analysis. Crop Sci. 21:355-358.

Shenk, J. S., K. H. Norris, R. F. Barnes and G. W. Fissel.
1977 Forage and feedstuff analysis with an infrared reflectance spectro-computer system. *In* Proc. 13th Int. Grassl. Cong., p. 454-463. Leipzig, German Democratic Republic.

Shenk, J. S., M. O. Westerhaus, and M. R. Hoover.
1979 Analysis of forages by infrared reflectance. J. Dairy Sci. 62:807-812.

Shenk, J. S., M. O. Westerhaus, and M. R. Hoover.
1978 Infrared reflectance analysis of forages. *In* Proc. Int. Grain and Forage Harvesting Cong., p. 242-244. Am. Soc. Agric. Eng., St. Joseph, MI.

Starr, C., A. G. Morgan, and D. B. Smith.
1981 An evaluation of near infrared reflectance analysis in some plant breeding programmes. J. Agric. Sci. (Camb.) 97:107-118.

Stermer, R. A., Y. Pomeranz, and R. J. McGinty.
1977 Infrared reflectance spectroscopy for estimation of moisture of whole grain. Cereal Chem. 54:345-351.

Templeton, W. C., Jr., J. S. Shenk, K. H. Norris, and others.
1983 Forage analysis with near-infrared reflectance spectroscopy—status and outline of national research project. *In* Proc. 14th Int. Grassl. Cong., p. 528-531. Westview Press, Boulder, CO.

Ward, R. G., J. D. Wallace, N. S. Urquhart, and J. S. Shenk.
1982 Estimates of intake and quality of grazed range forage by near infrared reflectance spectroscopy. J. Anim. Sci. 54:399-402.

Watson, C. A., D. Carville, E. Dikeman, and others.
1976 Evaluation of two infrared instruments for determining protein content of hard red winter wheat. Cereal Chem. 53:214-222.

White, C. H., G. A. Muck, M. Rearden, and P. Rotolo.
1978 The use of infrared reflectance analyzer for measurement of fat and total solids of dairy products. [abstr.] J. Dairy Sci. 61(1):120.

Williams, P. C.
1975 Application of near infrared reflectance spectroscopy to analysis of cereal grains and oilseeds. Cereal Chem. 52:561-576.

Winch, J. E., and H. Major.
1981 Predicting nitrogen and digestibility of forages using near infrared reflectance photometry. Can. J. Plant Sci. 61:45-51.

NIRS Instrumentation

K. H. Norris

Instrumentation Requirements for Quantitative Analysis by Diffuse Reflectance Spectroscopy

The near infrared region is generally defined as comprising the wavelengths from 700 to 3000 nm; however, most of the quantitative analyses by reflectance are done in the 1200- to 2500-nm region. Below 1200 nm, the absorption bands are so weak that quantitative measurements by reflectance are difficult; and above 2500 nm, the absorption bands are so strong that measurements are difficult. In addition, the signal-to-noise ratio falls off rapidly at the longer wavelengths because the energy from a tungsten lamp falls off rapidly. As a result, the most useful region is from 1200 to 2500 nm.

The NIRS method of analysis requires very high wavelength reproducibility, typically within a standard deviation of less than 0.02 nm. The wavelength accuracy requirements within one instrument are not very important, and a tolerance of ±5 nm is acceptable. However, if calibrations are to be transferred from one instrument to another, corrections must be made for the wavelength differences.

Diffuse reflectance (R) in the near infrared region will seldom be less than 1 percent and never greater than 100 percent. Expressed as log (1/R), the minimum value will be 0.0 and the maximum approximately 2.0. Therefore, an instrument should have good linearity over the photometric range from 0 to 2.0 in log (1/R), and a linearity of within 1 percent over this range is generally considered adequate. The accuracy of the photometric scale within one instrument is not important as long as it is stable. However, between-instrument comparisons require a photometric accuracy of better than 1 percent; otherwise, corrections must be applied. Very small changes in reflectance must be measured in the NIRS method; therefore, photometric noise is a very important consideration. Noise levels of 0.0001 in log (1/R) can limit the performance of a measurement.

Three different optical geometries are being used in commercial NIRS instruments. These are illumination at an angle of 0° and collection at an angle of 45° (as used in instruments manufactured by Pacific Scientific Corporation and Labor-Mim Company), illumination at an angle of 0° and collection of all reflected radiation with an integrating sphere (as used by Technicon Industrial Systems and Percon Corporation), and illumination at a small angle and collection at an angle of 0° (as used by Dickey-john Corporation). Each of these geometries has advantages and disadvantages, but they all provide adequate reflectance data for quantitative analyses.

Reflectance measurements require a reference standard. The National Bureau of Standards recommends the use of a commercial plastic, Halon, (Weidner and Hsia *1981*) as a reflectance standard for the ultraviolet, visible, and near infrared regions. A more rugged standard is desirable for the NIRS method, and a special ceramic has been adopted as the working standard. This ceramic has a uniform reflectance across the total near infrared region, as shown in figure 1. This standard can be cleaned with soap and water or with alcohol to remove contamination. Instruments with an integrating sphere use the interior coating of the sphere as a reference. The spheres are specially coated with gold to provide a uniform, high, and diffuse reflectance over the near infrared region. Pressed sulfur has been proposed as a reference standard (Tkachuk and Kuzina *1978*), but it has not been widely used despite its excellent reflectance properties in the near infrared region.

The reflectance of the reference must remain constant to maintain the calibration of an NIRS instrument. The absolute value of the reflectance is not of great consequence; but if calibrations are to be transferred from one instrument to another, any differences in the reference must be accounted for. It appears that the ceramic working standard is very stable and does not represent a signficant source of error within one instrument, but between-instrument variability occurs because the position of the ceramic and its backing affect the measured reflectance.

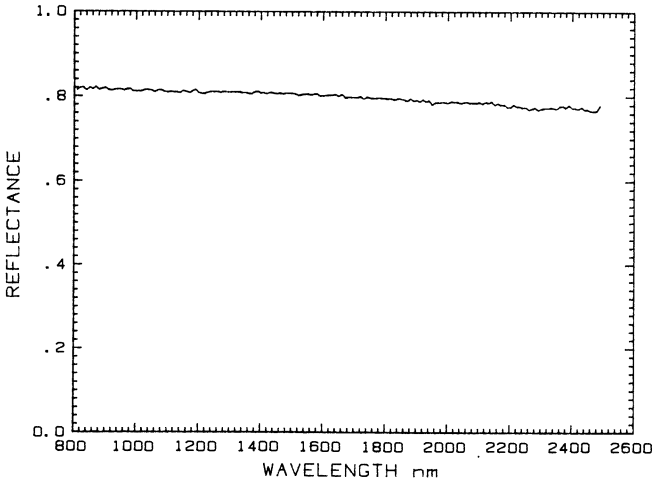


Figure 1.—Absolute reflectance spectrum for a ceramic reference standard. Data provided by National Bureau of Standards.

Wavelength-Scanning Instruments

Several groups have developed scanning monochromator instruments for their use in NIRS research (McClure and Hamid *1980*, Norris et al. *1976*, Shenk and Westerhaus *1977*). These instruments are various modifications of prism-grating spectrophotometers originally manufactured by Cary. The most common modification uses only the monochromator from the Cary and replaces the sample optical system with one that provides single-beam illumination at an angle of 0° and collection of the reflected light by a bank of four photocells at 45° (fig. 2). The sample geometry is essentially the one used by Pacific Scientific Corporation in several models.

The light beam is chopped into an alternating on-off beam, and synchronous detection is used with the photocell signal to reduce the effects of temperature change, amplifier drift, and stray radiation. Large-area, lead sulfide detectors are used to measure the radiation reflected from the sample. The signals from the detectors are summed into a high-gain, low-noise amplifier, and the amplified signal is converted to a digital signal with different types of signal sorting and analog-to-digital converters. The digital signal is coupled into a computer for processing and storage. A signal representing wavelength is coupled into the computer in synchrony with the reflectance signal so that a reflectance value is stored for each wavelength point. The wavelength interval between the data points can be set by the operator for anywhere from 0.1 to 10 nm, with 1.0 to 2.0 nm being the most common choice.

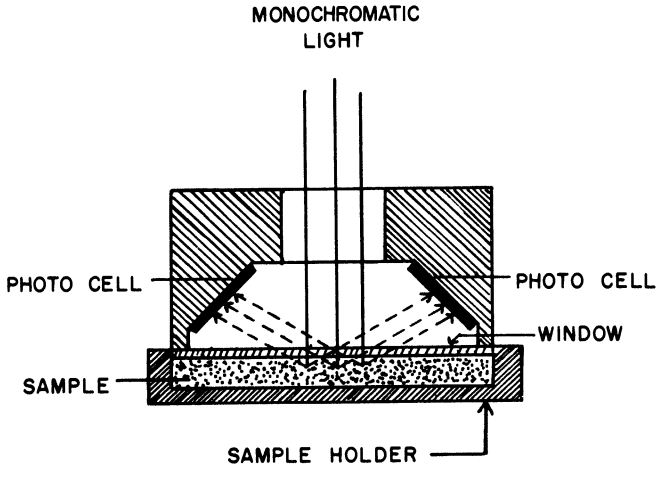


Figure 2.—Optical geometry for diffuse reflectance measurements based on using 0° illumination and 45° collection with large area photocells.

In operation, the spectrum for the reference standard is measured and stored in the computer. The spectrum for the sample is then measured, and the corrected sample reflectance spectrum is computed. The reflectance spectrum for each sample may be recorded on disk or magnetic tape for future analysis. The most common analyses use log (1/R); therefore, the data are often converted to log (1/R) before being stored on disk or tape.

The Cary monochromator has variable slits, so the spectral resolution can be varied. The low energy output of the monochromator limits the measurements to relatively wide pass bands, typically 5 to 10 nm or slit widths from 1.5 to 3.0 nm. The Cary monochromator provides excellent stray light rejection and excellent wavelength precision. It may be operated at scanning speeds as high as 10 nm per second, so the whole spectrum from 1000 to 2600 nm can be scanned in less than 3 minutes.

Two U.S. firms have produced and marketed scanning monochromator instruments specifically for rapid quantitative NIRS analyses. The model 6100 grating monochromator instrument was introduced by Neotec (now a subdivision of Pacific Scientific Corporation) in 1978 and has been marketed under several different model numbers, the major distinction being the computer used with each of the models. This instrument uses a large, high-efficiency grating to achieve high energy output and low noise with a rapid scan (Landa *1979*). The spectral region from 1100 to 2500 nm is scanned five times per second, and multiple scans are averaged to minimize the noise.

Spectral data may be collected at wavelength intervals of either 1.0 or 2.0 nm, with 2.0 nm being the most common. The instrument uses the sample geometry shown in figure 2 and operates in the single-beam mode. A ceramic reference sample is scanned, and the spectral data are stored in the computer, which uses the data to correct the reflectance of the sample. From 20 to 100 scans are averaged for both the reference and the sample: but since the scans are made five times per second, a low-noise, corrected-reflectance spectrum of a sample can be obtained in 20 seconds. The reflectance data are normally converted to log (1/R) before storage on the disk. The instrument is operated at a fixed slit width providing a nominal bandpass of 10 nm. An accessory is available for diffuse transmittance measurements of solutions and slurries. The instrument can also operate in the 600- to 1100-nm region if silicon detectors are used instead of the lead sulfide detectors.

The model 6100 monochromator was originally marketed with a Data General computer and a floppy disk. It now comes with a North Star Computer, including both a hard disk, a floppy disk, and a full line of operating software for collecting, plotting, and analyzing the spectral data. The software includes a selection of several regression procedures for calibration. The software also includes diagnostic routines to test the performance of the spectrophotometer.

The model 6100 monochromator has also been marketed with a Digital Equipment Company computer, and it is this version that was adopted for use by the USDAs National NIRS Forage Research Project Network. These models are typically equipped with a removable hard disk and dual floppies for storage of data and programs. The operating software for these units was developed as part of the Forage Network, and the backup support for this software comes from the Network. This software, which will be described in a later section, provides full capabilities for collecting and analyzing the spectral data. If also provides routines for testing instrument performance.

The model 500 grating monochromator instrument was introduced by Technicon Industrial Systems in 1982. Although this model also uses a large, high-efficiency grating to achieve high energy output, it is otherwise designed quite differently from the model 6100. Model 500 operates in a double-beam mode using an integrating sphere (fig. 3) to collect the reflected energy as well as to provide the dual-beam capability. The wavelength drive is powered by a stepping motor which enables both continuous scanning from 1100 to 2500 nm and rapid zeroing in on specific wavelengths or wavelength regions. The instrument requires 90 seconds for a complete scan when measurements are made at 2-nm intervals from 1100 to 2500 nm. During this scan, the reference and the sample are both measured, and the computer reads out the corrected reflectance or log (1/R) as desired. The monochromator slits are configured to provide a bandpass of 10 nm across the spectrum.

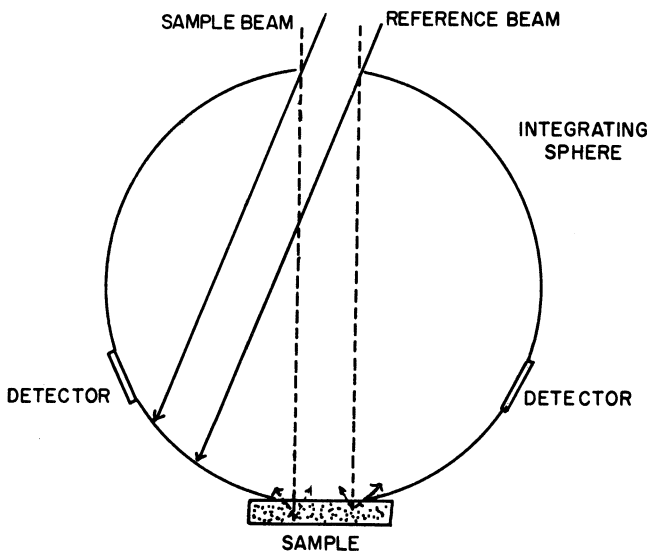


Figure 3.—Optical geometry for dual-beam diffuse reflectance measurements based on using an integrating sphere to collect the reflectance.

Model 500 is normally equipped with a Hewlett Packard computer and two floppy disk drives, but a hard disk is available. Software is provided for collecting and processing the spectral data, and several choices of regression procedures are included for calibration. The software also incudes diagnostic routines for testing instrument performance.

The model 500 monochrometer has also been marketed with a Digital Equipment computer; but in this configuration, the software support is from the Forage Network.

A scanning monochromator instrument has also been developed in Hungary by Labor-Min but has not been marketed in the United States. This instrument operates in a single-beam mode using the geometry shown in figure 2. It scans the spectral region from 1200 to 2400 nm in 60 seconds and has a nominal bandpass of 15 nm. It uses a small grating monochromator and does not provide the same signal-to-noise ratio as the larger grating instruments. The instrument incorporates a dedicated microcomputer to collect and process the reflectance data. Software is provided for limited data analysis.

L. T. Industries has announced a rapid scanning monochromator instrument for the near infrared region, but details of the instrument have not been made available.

Scanning-Filter (Tilting-Filter) Instruments

The transmitted pass band of an interference filter moves to shorter wavelengths as the filter is tilted away from perpendicular to the incident source. The effective wavelength of a filter may be shifted as much as 10 percent without serious distortion of the pass band. Therefore, it is possible to scan a limited wavelength region by tilting an interference filter. One of the original commercial NIRS instruments used this principle with three interference filters in a paddle wheel to scan the spectral region from 1800 to 2310 nm. The

principle is still used by Pacific Scientific Corporation in all of their filter-type NIRS instruments, that is, models 31, 41, 51, 101, 102, 4250, and 7000. These models used the optical geometry shown in figure 2 for reflectance measurements and incorporate various levels of computer capability to collect and process the reflectance data.

Tilting-filter instruments are generally marketed as calibrated instruments with stored calibration constants that allow the user to obtain digital readouts of from two to five constituents for each of several products. The calibration constants can be changed by the user to adjust to new products or new constituents. Some of these instruments can also be used in a transmittance mode; and by changing filters, it is possible to change the effective wavelength region. Silicon detectors replace the lead sulfide detectors for operation in the 600- to 1100-nm region.

These instruments may be used as fixed-wavelength instruments or as limited-range scanning spectrophotometers. As scanning spectrophotometers, they offer the capability of first or second derivation treatment of data, and they are most often used in this mode. The instruments all incorporate a ceramic reference standard which is measured between each sample, and the computer provides the output as log (1/R).

Fixed-Filter Instruments

Fixed-filter NIRS instruments have been developed by four firms. The first model was developed by Dickey-john Corporation and was marketed as the 2.5A Grain Analysis Computer by Technicon Industrial Systems. It is a six-filter instrument with wavelengths chosen, according to the specifications of USDA's Instrumentation Research Laboratory, to measure water (1940 nm), carbohydrates (2100 nm), protein (2180 nm), and oil (2310 nm) and to serve as references of minimum absorption (1680 and 2230 nm) for these constituents. The six narrow-band interference filters are mounted in a wheel so that the sample is successively illuminated for brief times with radiation from each of the filters. The optical geometry provides for illumination of the sample at a small angle and collection of the reflected energy at zero angle. A single lead sulfide detector is used to measure the radiation from the sample. The detector incorporates a thermoelectric element for cooling the detector to enhance the signal-to-noise ratio. The 2.5A model used an analog computer to sort the reflectance values and compute the constituent values. It has since been upgraded with a digital computer to GAC III models 640, 650, 660, and 800. The new units are marketed directly by Dickey-john Corporation. The newer units use the same six wavelengths originally chosen, with options of adding up to four additional wavelengths. The Dickey-john units have an additional feature that allows rotating the sample during measurement to minimize sample variability. All of these instruments operate in the single-beam mode, with a ceramic reference standard being measured between each sample.

Fixed-filter instruments for NIRS have also been developed by Technicon Industrial Systems under the model numbers of 300 and 400. These instruments use the integrating sphere geometry of figure 3 and operate in a double-beam mode, with the interior of the sphere being the reference standard. These units use the same six wavelengths that were originally chosen to measure oil, moisture, and protein, but the 400 offers up to 19 wavelengths. The standard choice for the 19 are 1445, 1680, 1722, 1734, 1759, 1778, 1818, 1940, 1982, 2100, 2139, 2180, 2190, 2208, 2230, 2270, 2310, 2336, and 2348 nm. These instruments incorporate a microcomputer to process the data into constituent analyses. A liquid-sample compartment is available for the Technicon 400 and 500 instruments, and it has been used for milk and dairy products as well as other liquid samples. The sample is filmed over a ceramic disk, and the radiation is transmitted through the liquid, reflected by the ceramic, and transmitted back through the liquid into the integrating sphere for measurement. This mode of operation is referred to as the "Transflectance Mode."

A fixed-filter instrument has also been developed in Germany by Percon. This instrument is very similar to the Technicon filter instruments, using the same integrating sphere geometry and double-beam mode. The major difference is the use of a flow-through sample compartment rather than a separately packed sample cup.

Instrument Error Sources

The major source of noise in NIRS measurements is the sample itself: sampling errors, sample instability, nonuniformity errors, sample-packing errors, errors in the laboratory analyses, or errors resulting from attempting to use NIRS for constituents having no near infrared signal to measure. This is true because the instrument manufacturers have done an excellent job of reducing the instrument noise to a minimum. The noise contribution of the instrument can be separated into long-term noise or drift and short-term noise. The major contributors to long-term drift are changes in ambient temperature and degradation of components. The instruments are designed to operate in a fairly constant temperature environment, so changes of 10 to 20 degrees will cause changes in the constituent readings. Component degradation may simply be due to an ac-

cumulation of dust on the reference standard or on windows, lenses, or other optical components; or it may be due to mechanical wear of moving parts or to deterioration of electronic components. Such changes can produce major errors or complete malfunction. Instruments operating in the double-beam mode are less subject to some of these errors because they affect both the reference and the sample, but errors resulting from dust on the sample window and errors from wavelength changes are not cancelled by the double-beam optics.

Wavelength instability is the main source of short-term noise in wavelength-scanning instruments. Detector and input amplifiers represent the next most common source of short-term noise for instruments which are functioning properly. Noise outside the instrument also can be a problem and may derive from brush-type motors, heater switches, and similar electrical devices which generate electrical pulses. Such pulses can be carried on the electrical lines into the instrument or can be transmitted as radiofrequency pulses through the air.

Single-beam instruments require very stable light sources. Therefore, the lamp and lamp power supply can be sources of noise. Dual-beam instruments must switch beams from reference to sample very fast, and this switching action can be a source of noise in such instruments.

Other sources of noise include stray light, photometric nonlinearity, and surface reflectance. All of these introduce nonlinearity errors into the measurement; but since the intensity of the reflectance signal is not linearly related to constituent content, any nonlinearity effects tend to be masked. These errors become a more serious problem for multiple-instrument applications in which it is desirable to transfer calibrations from one instrument to another.

Stray light is of two forms: room light reaching the detectors because of inadequate shielding and light leaking through the monochromator because it is not blocked by the dispersion element (grating or filters). Room light is seldom a source of difficulty if the instrument is operated properly (sample drawer closed, and so forth). The stray light in a well-designed grating monochromator is typically less than 0.1 percent. This would introduce an error of 10 percent in measuring an isolated absorption band whose log (1/R) value is 2.0. However, isolated absorption bands do not occur in forage samples, and the presence of the other absorbers reduces the stray-light effect. The maximum log (1/R) value for dry forage samples is less than 1.0; and at this level, the stray light of 0.1 percent causes an error of less than 1 percent. High-moisture samples such as fresh silage will have log (1/R) values exceeding 2.0; and for these samples, stray light becomes a more important source of error.

Well-designed filter instruments also should have less than 0.1 percent stray light, but an instrument with poor or damaged filters could have very high stray light. The combination of the spectral response of the lead sulfide detector and the spectral emission of the tungsten lamp results in a peak energy response in the 1300- to 1600-nm region. Therefore, stray light is most likely to affect measurements below 800 nm and above 2600 nm—the regions of the spectrum in which the available energy is very low.

Logarithmic amplifiers were used in the first instruments and were a source of photometric nonlinearity. But they are no longer used, and photometric nonlinearity is not a serious source of error on newer instruments. The nonlinearity generated from the surface reflectance signal represents a serious error for samples having high log (1/R) values. It is difficult to keep this reflectance signal below 0.3 percent, which represents an error of 3 percent for a sample having a log (1/R) of 1.0 and 33 percent for a sample having a log (1/R) of 2.0. Useful diffuse reflectance measurements can be obtained only for wavelengths at which the log (1/R) values are less than 2.0. Corrections for the surface reflectance can be included in the data collection to minimize this type of error, but such corrections are based on an average surface reflectance and are not specific to the sample being measured. Therefore, it has not been possible to adequately correct for the surface reflectance.

References

Landa, I.
1979 High-energy spectrophotometer for rapid constituent analysis in the range of 0.25-2.4 μm. Rev. Sci. Inst. 50(1):34-40.

McClure, W. F., and A. Hamid.
1980 Rapid NIR measurement of the chemical composition of foods and food products. Part 1 Hardware. Am. Lab. 12(9):57-59.

Norris, K. H., R. F. Barnes, J. E. Moore, and J. S. Shenk.
1976 Predicting forage quality by infrared reflectance spectroscopy. J. Anim. Sci 43(4):889-897.

Tkachuk, R., and F. D. Kuzina.
1978 Sulfur as a proposed near infrared reflectance standard. Appl. Opt. 17(17):2817-20.

Shenk, J. S., and M. O. Westerhaus.
1977 Infrared reflectance analysis of forages: grain and forage harvesting. *In* Proc. 1st Int. Grain and Forage Conf., Sept. 25-29, 1977, Ames, IA. Am. Soc. Agric. Eng., St. Joseph, MI.

Weidner, V. R., and J. J. Hsia.
1981 Reflection properties of pressed polytetrafluorethylene powder. J. Opt. Soc. Assoc. 71(7):856-862.

Computers

M.O. Westerhaus

Computers are vital to all phases of NIRS technology. During data collection, computers are used to control the operation of an NIRS instrument and to receive the NIRS measurements. Many instruments provide multiple readings which are averaged by the computer to improve precision. Sample measurements must be referenced to measurements of a standard and converted to log (1/R). These functions may be fulfilled by a microprocessor in the instrument or by an external computer.

Instrument calibrations also rely heavily on computers. NIRS data from many samples, along with their traditional quality determinations, must be stored in the computer's mass storage device. Elaborate computer programs read these files and derive equations relating NIRS data to sample-quality parameters.

Once the calibration equations have been derived, other programs are used to obtain quality determinations of similar samples based on their NIRS measurements. The sample analyses can then be output in a printed report or can be stored on disks for later use with farm management programs.

Computers communicate with the instruments by either a parallel or serial interface. Parallel interfaces transmit up to 16 data bits at a time using 16 separate lines. They therefore permit very fast transfer rates and are used when the instrument supplies data to the external computer at a very high rate. Serial interfaces transmit the 16 bits of data 1 at a time over 1 line, and this results in slower transfer rates. But because they are less expensive and their communication protocols more standardized, they should be utilized whenever possible.

In order to run all the NIRS related programs, microcomputers must meet certain hardware requirements. Computer memory must be large enough to hold the program plus all the NIRS data from a sample. Programs in external computers usually require 64K bytes of main memory; programs in internal computers often require less. Calibration programs can be designed to run quicker if more memory is available. Programs and data files must be stored on floppy disks or hard disks. Floppy disks are used for program and data interchange and backup. Hard disks are preferred for normal operations because they are more reliable, faster, and have a much larger capacity. The speed of the computer processor is critical only if the data are collected in real time at a fast rate. A high speed computer is helpful in the calibration process, which can take days on a typical microcomputer. Many computers can accept a float-

ing point processor which performs arithmetic on real numbers in hardware much faster than the standard processor. Currently, two scanning monochromators are commercially available for agricultural products. Pacific Scientific manufactures model 6350, a rapid scanning monochromator. Three sets of software can be used with it. Pacific Scientific has developed its software to run on Northstar Advantage and Data General computers. The software USDA developed for its research runs on Digital Equipment Corporation PDP 11 computers. Infrasoft International has modified the USDA software to run on IBM PC and Northstar Advantage computers. This software includes features helpful in business applications.

The second monochromator is the Technicon 500. The software supplied by Technicon runs on Hewlett-Packard computers. The USDA software has been modified to also interface to the Technicon 500.

Public Software

J. S. Shenk

As an outgrowth of the initial research with NIRS and forage analysis by Norris et al. (*1976*), the first spectrocomputer system dedicated to forage research was established at University Park, PA. A decision to use a DEC computer was made because it was the only computer company providing a service-and-maintenance contract in the University Park area. In addition, the PDP 11 computer family had provisions for enhancements, upgrading, and expansion.

At the time of this decision, software to collect, store, and process NIRS data on the DEC computer was not available. Developing such software was therefore the first task to be undertaken after the hardware had been put in place. One of the first programs to be written was for wavelength selection. A wavelength program written by Norris at Beltsville, MD, was obtained and then modified extensively to make wavelength selection automatic. This work was carried out by Wesley Mason, graduate student, and Rick Hoover, program analyst on the project (Shenk and Hoover *1976*).

At the same time, it became apparent that a series of support programs would be needed to manipulate files and data generated by the spectrocomputer. These programs became the nucleus for the public software currently in existence.

When the National NIRS Forage Research Project Network was in the planning stage, the need for this public software base became evident. As the Network plans began to develop, this software became a cornerstone of the Network structure. Several important provisions concerning the software were included to help the Network function efficiently.

1. New program concepts and evaluation research would be readily available to Network participants. This provision was needed because research on selection of math treatments and wavelengths was still in its infancy.

2. Statistical and NIRS terminology would be standard among Network participants. This would aid communication among Network participants.

3. The software would operate only on the PDP 11 computer equipment. This would standardize the computer and simplify training in and understanding of the computer portion of the analysis.

4. Dual floppy disks would be used for communication among locations. This medium would provide a standard mechanism for sharing information, spectra, and programs among Network participants.

5. The research could serve as a guide for private industry in their development of commercially available software.

The basic software-development philosophy was to provide the network participants with automated state-of-the-art programs that would have the highest degree of expert intelligence possible and still be user friendly. This was particularly true for the selection of math treatments and terms for the analysis equations.

It was apparent early in the initial research that plotting the spectra and using small sets of samples in calibration would accomplish little in advancing this technology in forage analysis. Every effort has been made to adhere to the initial programming concepts.

During the past 10 years of software development, many modifications and enhancements have been added to the programs. These can be easily recognized by comparing the present set of program options with the original programs. More important, however, are the features not seen by the user that provide protection in data collection, internal instrumentation diagnostics, fast yet accurate computation of the results, and rapid selection of equation terms in calibration. These are only a few of the new features that have made this public software package so successful (Shenk et al. *1981*).

In summary, the Forage Network provides three important benefits. First is the public software, enabling users to work cooperatively. Next is common hardware, enabling the transfer of data and new programs from system to system. Last, but not least, is the continued improvement of the software to provide the highest degree of accuracy in the easiest to use form.

The software consists of 14 programs written in FORTRAN IV. The programs were first written to run under RT-11 V3 on a DEC PDP 11/03 computer with dual RX01 single-density floppy disks. The programs have since been updated to RT-11 V4.0 and now run on DEC 11/23 computers with dual RX02 double-density floppy disks and RL01 and RL02 five-megabyte hard disks.

Main Programs

Data Collection Program (SCAN)

Two NIRS instruments can be used with this software: the Pacific Scientific 6350 and the Technicon 500. Each instrument collects different numbers of NIRS data points. The Pacific Scientific 6350 collects 700 and the Technicon 500 collects 750 NIRS data points. The program SCAN has the same format regardless of the instrument and has three modes of operation: (1) store NIRS data for each sample, (2) determine sample quality and store NIRS data, and (3) determine sample quality without storing NIRS data. Hundreds of determinations can be stored on disks for later statistical analyses. The program also generates reports on the line printer, based on the sample analysis.

Another feature of this program is that in the analysis mode the operator is warned when the NIRS data obtained on the sample differ dramatically from the NIRS data used to obtain the calibration equation. Finally, the program contains instrument diagnostics, such as instrument noise and wavelength accuracy and precision.

Data Manipulation Program (DATA)

This program allows the operator to add calibration data to a file of stored spectra for up to 10 laboratory or animal calibration variables. At any time thereafter the laboratory data and sample numbers may be listed and corrected, and additional variables may be added. The program also allows for the listing of NIRS data at specific wavelengths.

File Manipulation Program (FILE)

This program provides for splitting and combining files as well as for deleting samples and averaging together multiple NIRS scans of the same samples from a CAL or RAW file.

Calibration Program BEST

The objective of this program is to develop multiterm calibration equations. The program provides automatic evaluation of selected math treatments. The number of equation terms may vary from one to nine. Within-file analysis is provided, and samples can be omitted automatically from calibration and analysis. Disk reads are kept to a minimum. Single term division is provided, but the wavelength must be known by the operator. The program requires minutes rather than hours for a solution. Program terminology includes the following:

NIRS variables = NIRS math-treated data considered for inclusion in the calibration equation.

Term = NIRS variable used in the equation. Calibration equations with multiple terms take advantage of overtones and combination bands and adjust for interfering absorbers.

Differences = Math treatments similar to derivatives, where 0 = log (1/R), 1 = first differences, and 2 = second differences. First differences are the slopes and second differences are the changes in slopes. Differences provide for maximizing local information in the spectra and minimizing the effects of light scatter.

Division = Information at one portion of the spectrum divided by information from another portion of the spectrum. Only one division is possible, and the divisor wavelength must be known in advance by the operator. All wavelengths will be divided by the chosen wavelength.

Smoothing = A running average of NIRS variables after difference calculations.

Equation form = $y = b_0 + b_1 \cdot f_1(x_1) + b_2 \cdot f_1(x_2) \ldots$

Calibration Program CAL

The objective of this program also is to develop multiterm calibration equations. The program provides automatic evaluation of math treatments as well as fine tuning of the math treatments to maximize the information in a specific region of the spectra. Division math is also provided. Equation terms may vary from one to nine, and each term in the equation may have a different math treatment. Within-file analysis is provided as an aid to selecting the best NIRS variables for a calibration equation. The program is easy to operate when choosing the default options. Calculation time may be hours or days, depending on the computer or number of samples in the calibration file. Program terminology includes the following:

Fine tuning = Optimizing the math treatment to fit the information in a small region of the spectrum. This is accomplished by changing the distance between NIRS bands being subtracted and changing the segment length when smoothing. The program provides three default lengths (4, 16, and 24 nm) to be evaluated. The operator may choose these or up to nine lengths.

Division = Information at one portion of the spectrum divided by information from another portion of the spectrum. This division pair represents one term in the equation. Division may also be helpful in reducing the effects of light scatter.

Equation form = $y = b_0 + b_1 \cdot \frac{f_1(x_1)}{f_1(x_2)} + b_2 \cdot \frac{f_2(x_3)}{f_2(x_4)} \ldots$

Equation File Program (EQA)

This program provides for listing stored equations and creating new equation files from existing equation files.

Prediction Program (PRE)

This program provides for predicting sample quality from NIRS data stored in files. Analysis files can be listed; and data on samples can be deleted, added, and/or corrected.

Statistical Analysis Program (STAT)

This program provides for comparing laboratory versus NIRS analyzed data, NIRS analyzed versus NIRS analyzed data, or laboratory versus laboratory data in the same file. Statistics calculated include the mean, standard deviation of the variables, standard error of prediction, and r^2.

Equation Transfer Program (TRNSFR)

This program makes possible the transfer of a calibration equation from one instrument to another. Three files are necessary to accomplish the transfer: (1) the spectra file of samples collected on the first instrument, (2) the equation file from the first instrument, and (3) the spectra file for the same samples collected on the second instrument. The output will be the transferred equation file ready to use on the second instrument.

Evaluation Program (EVALU8)

The EVALU8 program compares and evaluates the results of the TRNSFR program. Using the equation and sample analysis from one instrument and the transferred equation and sample analysis from another, it generates comparative statistics. These include mean, standard deviation, *N*, bias, standard error of the difference (SED), standard error of the difference corrected for bias (SED(C)), r^2, standard error of calibration (SEC), and the average of SED and SEC, (AVE). A statistic summary which selects the best equation for each variable is printed also.

Spectral Transfer Program (NIRTRN)

This program transfers the spectra from one scanning monochromator to another or from the same monochromator to itself after major repair. The program makes a wavelength-by-wavelength transfer, using a set of transfer samples analyzed by the two instruments.

Auxiliary Programs

NEWVAR

This program allows the operator to generate up to five new variables from mathematical equations to a CAL or PRE file.

CHEM

This program allows the operator to copy chemical data from a CAL file, containing spectra and calibration data, onto a RAW file, containing only spectra.

ORDER

The purpose of this program is to order a file from the lowest number to the highest number for a specific chemical constituent.

PLOT

This program will graph many different types of information associated with the NIR spectra on a Tektronix graphics terminal. Spectral curves may be plotted, overlaid, and/or subtracted. Correlations between chemical data and NIRS wavelengths as well as scatter plots between two selected wavelengths or chemical data may be plotted. Cursor identification of wavelengths and overlay of plots from different files can be accomplished.

References

Norris, K. H., R. F. Barnes, J. E. Moore, and J. S. Shenk.
1976 Predicting forage quality by infrared reflectance spectroscopy. Crop Sci. 23:94-96.

Shenk, J. S., and M. R. Hoover.
1976 Infrared reflectance spectro-computer design and application. Advances in automated analysis. Tech. In. Cong. 2:122-125.

Shenk, J. S., I. Landa, M. R. Hoover, and M. O. Westerhaus.
1981 Description and evaluation of a near infrared reflectance spectro-computer for forage and grain analysis. Crop Sci. 21:355-358.

Shenk, J. S., M. O. Westerhaus, and M. R. Hoover.
1978 Infrared reflectance analysis of forages. *In* Proc. Int. Grain and Forage Harvesting Conf., p. 242-244, 252, Am. Soc. Agric. Eng., St. Joseph, MI.

Sampling

S. M. Abrams

An accurate measurement of forage quality is not possible unless the forage analyzed in the laboratory closely represents the mean of the population from which it originated. Poor sampling technique and inadequate numbers of subsamples are the largest sources of error in any forage analysis system.

Variation in population sampling exists at each sampling site and between sampling sites. Sampling a load of baled hay exemplifies these sources of variability. Variation between bales (between-site variation) occurs because different areas of the load came from different parts of the field or perhaps different fields. Additionally, one section of the load may have been exposed to more moisture or light than another section. Thus, sampling from only one bale in the load may result in an analysis that grossly overestimates or underestimates the quality of the entire load. Variation within the bale (within-site variation) can originate from several sources. The outside of the bale may have been exposed to weathering or oxidative processes. Depending on moisture content, the interior may have been subjected to oxidation or fungal growth. In either event, sampling from the surface will certainly give erroneous results. Another source of within-site variation is due to the physical qualities of hay. It is not uniform, consisting of leaves, stems, and contaminating plants. Careless sampling from the bale will often result in a sample with a disproportionate quantity of stems, making the hay appear poorer than it really is.

Hand sampling will generally not yield representative samples from baled hay, and a coring device (for example, the Penn State Hay Corer) is essential. This device consists of a stainless steel tube with a set of removable boring teeth at the sampling end and a fitting at the other end. Two types of fittings are available. One accepts an electric drill, and the other accepts a manual brace. The latter, due to the lack of power, is more difficult to use, particularly when sampling tightly packed bales.

One core taken from each of 20 randomly selected bales will adequately represent one lot of hay. A lot of hay is defined as having the following characteristics: It derives from the same field and from the same cutting, is cut within a 48-hour period, and is uniform in maturity. The corer should be inserted into the bale from the long end and the sample extracted from the center. After sampling, the corer is emptied into a sample collection bag, using the wooden rod provided for this purpose. A check should be made to see that sufficient sample has been collected. Occasionally, with loosely packed bales, the corer will collect little hay, hitting mostly air pockets within the bale. When this occurs, the bale should be resampled from a slightly different position. Collected samples should be composited in one bag.

The composite sample should be mixed in the sample bag. Although mixing can be accomplished by hand stirring, the preferable method is sample rotation. The sample is placed in a plastic bag that is considerably larger than the collected sample, and the top is partially closed to leave a small hole. Then, the bag is blown up to its full dimensions, grasped shut, and rotated both vertically and horizontally to fully distribute the sample. It is preferable that the entire sample be sent to the testing laboratory. If the sample is too large, subsamples may be collected randomly from different areas of the bag. Analysis should take place as soon as possible after sampling.

Sampling silage is more difficult than sampling hay. Irrespective of silo type, only a portion of the silage mass is accessible. Due to the high moisture content, coring devices often do not operate satisfactorily; and silage is subject to rapid deterioration when removed from the anaerobic environment of the silo. For these reasons, estimates of silage quality are often less reliable than those of hay.

Only the face of the silage mass is available for sampling from bunker silos. Samples should be taken from six locations, two each from near the top, center, and bottom. Surface material should not be excluded, since it is also being fed. An even smaller amount of silage is available for sampling in an upright silo. Usually, the material is only accessible through an access door or a bottom unloader. The optimum time for sampling is at feeding, as the material is being unloaded from the silo. Random samples, no less than six, should be taken by hand at this time. If sampling at the door is necessary, the sample should be taken from as deep within the silage as possible, since silage at the door will often have been exposed to oxygen.

Silage is not as easily mixed as hay. If subsampling is necessary, silage samples should be mixed thoroughly by hand, reaching down to the bottom of the bag to prevent the fines from collecting at the bottom. It is preferable that the entire sample be sent for analysis. Samples should be dried at 60° C (140° F) for at least 48 hours before they are sent to the laboratory.

Sample Preparation

S. M. Abrams

Consistency of technique is crucial to the successful use of NIRS for feed analysis. Since NIRS is responsive to both chemical and physical properties of the sample, spectra will vary with the methodology used for sample drying or grinding. Differences in these procedures can easily lead to spurious results. An example of this problem was reported by Marum et al. (*1979*), who found that calibrations from forage samples dried in paper bags could not be used to analyze samples dried in cloth bags.

Moist samples (less than 88 percent dry matter) should be dried in a forced air oven at 60° C (140° F) prior to grinding. This temperature is sufficiently high to permit most of the water to be driven out and low enough not to significantly alter sample chemistry. After drying, the samples should be allowed to equilibrate with atmospheric temperature and humidity. Wet samples (for example, silage) should be dried a minimum of 48 hours.

An alternative drying method is to use a microwave oven, which increases speed of analysis. Drying times are reduced from over a day to several minutes. However, careless use of microwave ovens can result in charred samples due to overexposure to microwave radiation. Samples should be dried to a range of 90 to 94 percent dry matter. Trial and error is an important part of using the micowave oven, permitting the operator to judge the appropriate duration of drying needed.

Samples for NIRS analysis should be ground with a cyclone mill fitted with a 1-mm screen. Unpublished studies have demonstrated that precision of analysis is greater with samples ground in the cyclone mill than in a conventional Wiley mill because of the increased uniformity of particle size obtained. The mill should be blown out between samples to minimize cross-contamination, and the grinding ring periodically changed to ensure consistency of particle size across time.

It is often desirable to store samples for long periods of time in case recalibration is necessary or to provide samples to others for calibration. It is essential that these samples remain chemically unaltered. Changes in spectra can occur due to changes in moisture in the sample or due to the degradative action of aerobic or anaerobic micro-organisms. A sealed container is necessary to prevent moisture changes, and samples to be stored should be greater than 88 percent dry matter in order to limit microbial activity.

The storage container of choice is the retort pouch, constructed of aluminim foil sandwiched between two layers of plastic. The pouch is sealed on three sides and, after filling with sample, is sealed on the final side with a heat sealer. Care must be taken that the edges of the sealing surface are not crimped during sealing, as this will provide an avenue for moisture to enter or escape the bag. When a sample is needed, the end can be cut off with a paper cutter, a portion of the stored sample removed, and the pouch resealed.

Reference

Marum, P., A. W. Hovin, G. C. Marten, and J. S. Shenk.
1979 Genetic variability for cell wall constituents and associated quality traits in reed canarygrass. Crop Sci. 19:355-360.

Sample Handling
S. M. Abrams

The sample holder used by the Network is of a different design than the holders produced by the instrument manufacturers. The window is GE-124 quartz, 1.5 inches in diameter, and 0.045 ± 0.005 inches thick. Sample capacity ranges from a minimum of 0.75 g to a maximum of 1.75 g.

The sample is held in the holder with a disposable back made of rubber of foam core. The foam core back is cut from 1/8-inch foam core board and is approximately 0.002 inches larger than the interior of the holder. The paper back on the foam core is useful for recording sample numbers or other information. This back can be reused five or more times or until it no longer fits tightly. For removal, the foam core back is pried out of the sample holder with a special sharp-pronged tool. Care should be exercised so that the window is not scratched during removal.

When a back no longer fits tightly in a sample holder, its use can be extended by pinching the forage against the edge when inserting the back. The back should be discarded when it cannot be held firmly in the holder. The sample should always be tightly in place against the window.

The sample holder is usually cleaned with a camel hair brush or vacuum. If additional cleaning is needed, a soft tissue or lint-free cloth may be used. Finger prints and foreign material on the glass will cause analytical errors.

Maintenance problems with the sample holder primarily concern the window. If the window becomes loose or broken, it must be replaced and reglued into the holder ring. Care should be taken to use the same glue that was previously used by the manufacturer of the holder. Thirty-minute epoxy works satisfactorily. All excess glue must be removed from the window with acetone before using the holder for analysis.

Correct sampling procedures are important and the following method should be used. Open the storage container and stir the sample carefully with a curved spatula. Make sure the sample is well mixed, and take a random sample from the container. Continue to take random samples until the sample holder is level full. Take the disposal back and press it into the holder. Continue to press the back in until it is tight and level. As a check, turn the holder over and make certain the sample is firmly pressed against the window. If any abnormality is apparent, remove the back and repeat the procedure.

Instrument Operation
M. O. Westerhaus

NIRS measurements are very sensitive to the temperature of the instrument and the sample. Because of this, all instruments have controls to regulate internal temperature. These controls work well within normal room temperatures of 25 ± 5° C. The sample temperature, however, is not regulated by the instrument so the operator must make sure that the samples are at room temperature before they are analyzed. A constant relative humidity of 60 ± 2 percent is recommended.

If the room environment cannot be maintained within specifications, an alternate method must be employed in the calibration of the instrument so that the calibration will be useful over a range of environmental conditions. Spectra should be collected of the same samples under varying room environments. These samples will have the same chemical values; so when wavelengths are selected by the computer in calibration, the effects of changing room temperature and humidity will be minimized. This technique would also apply to analyses of samples that have different temperatures.

Instrument warmup is important, and the time required for it relates to the instrument model. It is recommended that most instruments be run continuously. The electronic components work best if they are operated continuously, and bulb life is now approximately 3 years. However, if the instrument is turned on from a cold condition, warmup time should be no less than 15 minutes and may require as much as 1 hour for some instruments.

Warmup time is especially important when using the National Forage Network software. The lamp in the Pacific Scientific 6350 monochromator is turned on by the SCAN program. The warmup is accomplished by connecting the instrument to AC current and typing in the word SCAN on the computer to activate the warmup period. The warmup sequence needed for other instruments may be found in the manufacturer's manual.

Once the instrument has been properly warmed up, there are two diagnostic tests which should be performed before beginning each work session. First, the noise level of the instrument needs to be ascertained. Second, the wavelength accuracy needs to be verified. Other diagnostic tests which can be performed on the monochromators are optimum scaling of the data and communications between instrument and computer.

The instrument noise level is described by two statistics. The calculations can be based on the ceramic referenced to itself or on the differences between two normal scans (sample referenced to the ceramic). In the latter case, the differences are multiplied by $\sqrt{2}$ to be comparable to the first case. In both cases, the log (1/R) values should be near zero. Deviations from zero are expressed as the average deviation (bias); and root mean square (RMS), expressed as log $(1/R)/10^6$.

The bias measurement characterizes any systematic change in log (1/R) level of scans taken over time. The absolute value of these numbers should usually be less than the RMS values. Larger values may indicate that the room temperature is fluctuating by more than 2° C. About half of these numbers should be positive and the rest negative. Bias values that are all positive or all negative indicate that a problem exists with the instrument and can affect the accuracy of the analyses.

RMS noise is the best single diagnostic of a potential instrument problem. The RMS values can range from a low of 10 to a high of 50 without affecting the analysis of most forage and feed products. The major effect of higher noise will be a lower level of analysis repeatability. In monochromators manufactured since 1983, the average noise level of 100 scans should be below 30. Only when these specifications on bias and RMS noise are not met should other diagnostic efforts be made.

Wavelength repeatability and accuracy are measured with a clear polystyrene petri dish. A scan of the polystyrene is made by placing the petri dish in the light beam and pulling out the sample drawer to expose the ceramic standard. This scan is referenced to measurements of the ceramic without the petri dish. The major polystyrene peaks are located and compared with the known locations at 1680.3, 2164.9, and 2304.2 nm. The standard deviation of repeated polystyrene peak measurements should be less than 0.05 nm. The average observed locations should differ from the known locations by less than 0.5 nm. Larger values usually indicate mechanical problems in the monochromator.

Since dust will accumulate in an instrument from a number of sources, it is important to implement a program of routine maintenance which will minimize dust contamination. The areas that must be kept clean are the ceramic standard, all parts of the drawer assembly, and the windows above and below the detectors. The dust filter in the instrument must be periodically cleaned. These areas should be cleaned with vacuum, brush, or soft tissue whenever dust can be seen to accumulate. Cleaning of the ceramic should be done at least once a week.

Directions for changing a lamp may be found in the instrument manuals. If the lamp is properly installed there will be no effects on a calibration.

Equation Selection

J. S. Shenk

Two basic types of analysis equations used in NIRS are narrow based and broad based. Narrow-based equations developed for finite populations have limited value beyond their use for those populations. Broad-based equations for infinite populations are more difficult to develop but have many uses. Examples of this latter type would be service equations to analyze research forages samples from a number of different scientists at a public laboratory or equations developed by private laboratories to analyze routine samples from the agricultural community.

Equations developed for infinite populations can be called universal equations. The word "universal" denotes that the calibration equation can be used for a rather large group of samples. An example would be a universal hay equation developed with samples from a wide range of forage species and mixtures of species for a large geographical area. Much of the disappointment that has been expressed in the use of NIRS is due to the lack of understanding of the universal equation concept.

Calibration is covered in detail in a later section of this handbook, but some insight into the concept may be helpful at this point. Samples selected for calibration must contain all the variables affecting the chemical and physical properties of the samples to be analyzed. Accuracy of the equation is directly related to this requirement. The data in table 1 demonstrate some of the characteristics of narrow-based equations and universal equations. The data were obtained for hay samples from a factorial field study designed to test the universal equation concept (Templeton et al. *1983*). Samples were from different species, years, harvests, and methods of drying. The factors in the experiment were as follows:

Factor		*Particulars*
Species	(5)	Alfalfa, red clover, birdsfoot trefoil, orchardgrass, timothy.
Years	(2)	1979 and 1980.
Harvests	(3)	First, second, and third.
Oven drying	(3)	At 50, 65 and 80° C.
Freeze drying	(1)	Immediately after harvest.
Field drying	(3)	1, 2, and 3 days after harvest.

The results in table 1 demonstrate that equations developed for analyzing a single species, such as orchardgrass, did not provide accurate analysis when used for a different species, such as red clover. Likewise, equations developed from samples that were oven dried were not very accurate when used to analyze field-dried samples. However, the universal equation derived with one half the samples, representing all sources of chemical and physical variation, had an acceptable level of accuracy over all samples in the experiment.

Table 1.—*Standard errors of prediction (%) associated with a single value of 5 forage-quality parameters predicted by NIRS*

Equation derived using—	Equation applied to—	Quality parameters				
		CP[a]	NDF[b]	ADF[c]	Lignin	IVDMD[d]
Birdsfoot trefoil	Alfalfa	0.40	1.75	2.93	0.69	2.14
Orchardgrass	Red clover	1.67	15.90	4.24	2.08	3.88
Orchardgrass	Timothy	.40	1.35	1.57	.54	2.28
Oven-drying at 65° C	Oven-drying at 75° C	.50	1.62	.93	.62	2.23
Oven-drying at 65° C	Field-drying for 3 days	.99	4.46	3.19	1.34	5.69
Harvest 1	Harvest 2	.94	2.52	1.52	1.42	3.63
Half of samples	Other half	.66	1.97	1.02	.59	1.80

[a] CP = crude protein.
[b] NDF = neutral detergent fiber.
[c] ADF = acid detergent fiber.
[d] IVDMD = in vitro dry matter digestibility.

Universal equations will be extremely important to future use of the NIRS technology. These equations will be needed so that precalibrated instruments can be made available both to the public and to private industry. With experience it is possible to collect a set of samples with universal characteristics without using the factorial approach. The key is to remember the simple theory behind the calibration concept.

Calibration equations may be obtained from the instrument owner or may be purchased from public or private sources. If the equation has been developed for the instrument by the owner, the remaining sections of this handbook should be reviewed and understood. If the analysis equation is to be obtained from a public or private source, the user must make sure that—

1. The equation is appropriate for the instrument, the computer software and instrument are comparable, and wavelengths and math treatments in the equation match the hardware configuration of the instrument.

2. An acceptable procedure is available for transferring the analysis equation to the instrument. (Handbook section on transfer technique should be read.)

3. The analysis equation has been developed specifically for the target application.

4. The universality of the equation covers at least 95 percent of the samples to be analyzed.

5. The equation contains all of the analyses needed.

6. There is no doubt about the type and reliability of the chemical analysis used in the calibration procedure.

7. A demonstrated level of accuracy and precision is guaranteed with the equation.

8. The method of sample preparation is known and was taken into account when the calibration procedure was set up.

Reference

Templeton, W. C., Jr., J. S. Shenk, K. H. Norris, and others.
1983 Forage analysis with near-infrared reflectance spectroscopy: status and outline of national research project. *In* J. Allen Smith and Virgil W. Hays, eds., Proc. 14th Int. Grassl. Cong., p. 528–531. Westview Press, Boulder, CO.

Monitoring Analysis Results

J. S. Shenk

One goal of NIRS is to provide accurate and precise determinations of forage quality. To attain this goal, all components of the NIRS analysis must be functioning properly. Since several of these components can malfunction at any time, NIRS results must be continually validated.

Validation of NIRS analyses consists of comparing the NIRS determinations with those made by the reference procedure used to derive the calibration equation. Control charts are often used to display the differences between methods. Reference procedures include chemical analyses, animal performance measurements, and calibrations on a master NIRS instrument. A successful validation indicates that both the NIRS instrument and the reference procedure are operating properly. It is extremely important, therefore, that the reference method be checked for accuracy, precision, and stability.

To supplement validation of the final NIRS determinations, checks can be made on several of the NIRS analysis components. These include the instrument performance and appropriateness of the calibration equation. Monitoring each part of the analysis will provide the operator with the confidence that the analysis is under control or will help the operator identify any problems.

Instrument performance should be monitored by using the program options in SCAN to determine instrument precision and wavelength accuracy. (See "Public Software.")

Two check samples should be maintained in sealed holders and analyzed routinely. Since there is currently no way to preserve agricultural products perfectly, some minor changes are inevitable. Large changes in sample analyses indicate either that the instrument has changed or that the check samples are no longer properly sealed. Similar changes in sample analyses indicate a problem in the instrument; dissimilar changes in sample analyses indicate a problem with one of the sealed samples. Caution should be exercised in the interpretation of these changes until some experience is gained with the effective lifetime of the product in the sealed holder.

Using the *H* statistic is the best way to validate an NIRS analysis without further analysis by the reference method. Determinations with large *H* values are flagged on the computer printout with asterisks. The critical *H* level can be set by the operator; but, normally, "*" represents three times the expected (average) value of

H and "**" represents four times the expected value of H on the file of calibration samples. An occasional determination with a large H value is expected, but if more than 20 percent of the determinations for a single sample have large H values, the analysis should be reported with caution. Another test is to analyze a number of similar samples and obtain the average H value with the program PRE. In general the average should be no larger than three times the expected value.

Another evaluation may be made by checking if the determinations fall into a normal range for the constituent. The operator should be aware that determinations with large H values often lie outside the normal range. Samples with determinations that fall outside the normal range but do not have extreme H values should be further analyzed by the reference method.

Two validation procedures are available for the two general operating situations in NIRS use. In the research situation, where large groups of similar samples are run at one time, a subgroup of samples should be randomly selected and analyzed by the reference procedure. The differences between the two analyses should be checked for excessive bias and variability. In the routine laboratory situation, where the instrument is used to analyze a wide variety of samples arriving at different times, a quality control procedure should be set up to check for time-related shifts in the population. At regular intervals a minimum of three samples of each product should be selected and analyzed by both NIRS and the reference procedure. The mean and standard deviation of the differences should be plotted over time and examined for any nonrandom pattern.

Two types of control tests are needed. The first test detects shifts in the average values, and the second test detects the magnitude of the difference between the reference procedure and the NIRS analysis. A control table for the research environment and a control chart for the routine analysis laboratory can be developed.

To develop these control tables or charts, control limits must be established. The method for calculating these limits is as follows:

1. Calculate the mean difference between the NIRS and reference procedures for the control samples. The mean control limits (MCL) are calculated from the equation

$$MCL = \frac{t \cdot SEC}{\sqrt{N}}$$

where t is the appropriate value ($\alpha = .05$) from the t table, SEC is the standard error of calibration for the analysis equation, N is the number of samples in the control set.

2. Calculate standard deviation (SD) for the control samples. The SD control limits (SDCL) are calculated from this equation

$$SDCL = \sqrt{\frac{x^2}{N-1}} \cdot SEC$$

where x^2 is the appropriate chi square statistic ($\alpha = 0.05$) and SEC is the standard error of calibration for the analysis samples.

3. If the control test is being conducted between master and slave NIRS instruments, substitute the SED between instruments for SEC in the MCL calculation and substitute the standard error of a difference corrected for bias (SED(C)) for SEC in the SDCL calculation.

4. Choose the t and x^2 values so that the control limits include 99.9 percent of the observations under control. Data on such a chart should not fall outside the control limits and should not exhibit any systematic trends. If control samples fall outside the limits, and all other instrument and sample processing checks are in order, recalibration may be necessary for routine analysis. In a research situation, the remaining samples should not be analyzed by the equation.

This monitoring system can provide a means of attaining the full potential of routine NIRS analyses.

Introduction

F. E. Barton II

This section of the manual and appendixes 2 and 3 (detailed forage quality considerations) will treat, in some detail, the laboratory procedures for analyzing forage samples and for using those data to develop, calibrate, and validate equations for predicting forage constituents by NIRS. The section is divided into three parts: (1) NIR spectra of forage samples, (2) animal response prediction, and (3) verification of NIRS results by chemical analysis. The material in this section is meant to amplify the NIRS procedure by discussing the chemistry of the forage sample, along with its spectral characteristics and its structure. While we can be quite detailed, we cannot be comprehensive as to all the analytical principles and laboratory techniques required. The reader is encouraged to make use of the specific bibliography at the end of each report or appendix and to consult the general references listed below for additional background. Additionally, it is recommended that a good statistics text be used as reference. Since statistic methods vary and some particular methods may at one time be more appropriate than others, no general reference is given. The specialized statistical treatment of NIR data is given in another section of this manual.

References

Analytical Methodology

AOAC.
1980 Methods of Analysis. 13th ed. Association of Official Analytical Chemists, Arlington, VA.

Barnes, R. F.
1973 Laboratory methods of evaluating feeding value of herbage. *In* G. W. Butler, R. W. Bailey, eds., Chemistry and biochemistry of herbage. Academic Press Inc. Ltd., London.

Instrumental Analysis

Willard, H. H., L. L. Merritt, and J. A. Dean.
1974 Instrumental methods of analysis. D. Van Nostrand Company, New York.

Forage and Food Evaluation

Mohsenin, N. N.
1984 Electromagnetic radiation properties of foods and agricultural products. Gordon and Breach Science Publishers, Inc., New York.

Wheeler, J. L. R. D. Mochrie.
1981 Forage evaluation: concepts and techniques. American Forage and Grassland Council, Lexington, KY, and CSIRO, East Melbourne, Victoria, Australia.

Martens, H., and H. Russwurm, Jr.
1983 Food research and data analysis. Applied Science Publishers, Ltd. Barking, Essex, IG11 OSA, England.

Chemometrics

Kowalski, B. R.
1977 Chemometrics: theory and applications. ACS Symposium Series 52. The American Chemical Society, Washington, DC.

Spectra

F. E. Barton II

The near infrared (NIR) region of the electromagnetic spectrum lies between the visible (VIS) and mid-IR regions. It is defined primarily by the signal-to-noise (S/N) response of the material used for detectors in the region. The energy response of a photomultiplier tube falls off at wavelengths above 600 nm, and lead sulfide is used for a detector in the NIR. It is generally accepted that the region of the electromagentic spectrum between 800 and 2500 nm is the NIR (Willard et al. *1974*). The spectra of forages in the region consist of harmonic overtones of mid-IR fundamental bands and combination bands generally associated with hydrogen-atom stretching-bond deformations. In general, the overtones occur from 800 to 2100 nm, and the combination bands from 2000 to 2500 nm. An overtone is defined as a harmonic (that is, one-half, one-fourth, and so on) of the frequency of the mid-IR fundamental absorption band. For example, the O-H and N-H stretching frequencies at 1400 and 1500 nm, respectively, are the first overtone (one-half the frequency) of the same bands in the mid-IR (2800 and 3000 nm, respectively). A combination band is the sum of difference between the frequencies of two or more fundamental or harmonic vibrations. In the NIR, the combination bands are usually harmonics.

The absorption frequencies of the various functional groups in the IR are generally given in reciprocal centimeters (cm^{-1}). The wavelength designation used in the NIR is either micrometers (μm), for example, 2.178 μm, or nanometers (nm) for example, 2178 nm. These three wavelength designations are interconvertible through the following equation:

$$\frac{1 \times 10^4}{\mu m} = cm^{-1}, \text{ where } 1\ \mu m = 10^3 nm$$

Theoretically, the NIR spectrum of any functional group can be calculated if the following were known: Its symmetry number, the wavelength of its fundamental vibrations, and the possible combinations and harmonics. Such calculations are possible for diatomic and triatomic systems but are not practical for complex biological samples such as grass samples. First approximations of absorption band assignment can be made by referencing to Colthup tables in the IR as in Willard et al. (*1974*, p. 172-173) and in the NIR as in Kaye (*1954*). The two part review of NIR spectroscopy by Kaye (*1954*, *1955*) is the most authoritative work to date. In it are discussed the identification of spectral features, instrumentations needs, techniques, and analytical applications. Although instruments have since been improved, the information provided is still useful because it is basic. The review should be read in light of the fact that the samples were in solution and their spectra obtained in the transmittance mode. Present NIR instruments for forage analysis operate in the diffuse reflectance mode, and the samples are solids.

NIR spectra arise from the heat-induced asymmetric stretching vibrations of hydrogen bonds in the functional groups of molecules. The individual stretching modes can be visualized as in figure 6-1 and 6-2 in Willard et al. (*1974*, p. 151).

NIR analyses offer some unique advantages. First, the S/N ratio is quite good. While the intensities of the overtone and combination bands are orders of magnitude weaker than those of their corresponding fundamental bands, this disparity is not a disadvantage because the light sources used are intense and the lead sulfide detector is quiet. This results in a good S/N ratio. The low molar absorptivities (small extinction coefficients) yield uniform signal intensities, so dynamic range can be easily handled. The log (1/R) spectra of grasses (fig. 4*A*) are broad, with few well defined features; so they can easily be, digitized with a reasonable number of data points. Each of the humps and curves present in a log (1/R) spectrum is a composite representing several unresolved vibrations. In almost all instances, the individual vibrations are related, and that relationship is repeated several times across the spectrum. Considerable intercorrelation is therefore evident in the NIR, and allows multiterm calibration equations to be developed for relating the complex chemical composition to the equally complex spectrum. The caution to be exercised is that intercorrelation can lead to overfit if too many terms are used.

Norris et al. (*1976*) reported that log (1/R) and its second derivative are the two most promising mathematical treatments for forages. The log (1/R) spectrum is analogous to an absorbance spectrum. The second derivative can be obtained by successive derivatization over typically 20 to 30 nm with some smoothing or by using three equal spectral lengths, (*A*, *B*, and *C*) of 20 to 30 nm and the following equation:

$$\frac{A + C - 2B}{2}$$

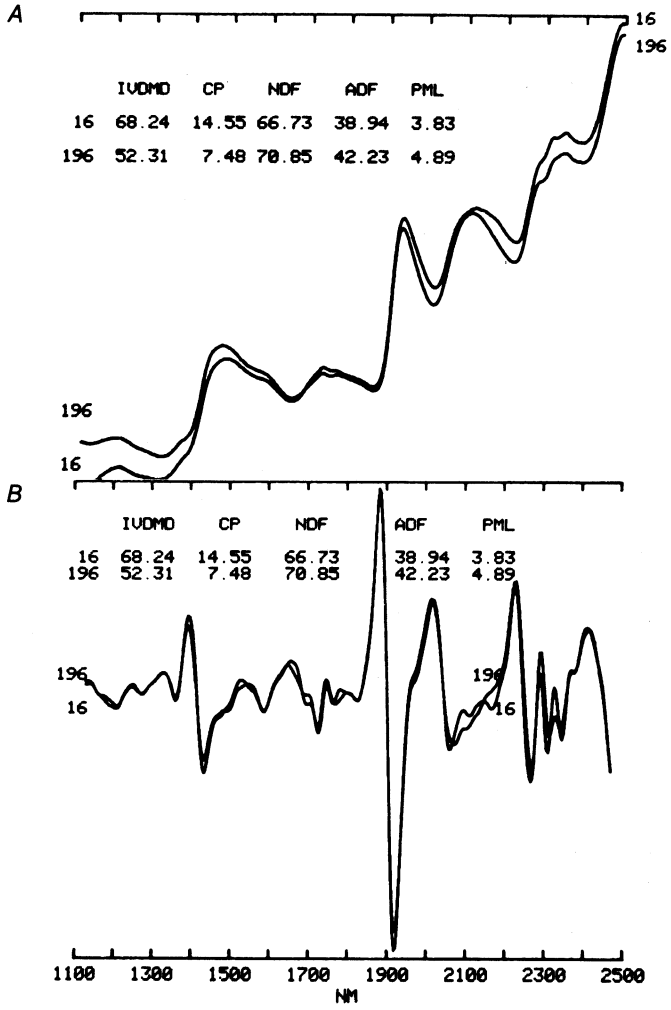


Figure 4.—NIR spectra of high quality (sample 16) and low quality (sample 196) Old World bluestem grass: *A.* log (1/R), *B.* second derivative of log (1/R).

Maddams (*1980*) reviewed curve fitting and derivative techniques to show both their utility and their limitations. Most of the forage work has been done using second derivative techniques to show both their utility and limitations. This mathematical treatment generally gives better resolved spectra, with a minimum occurring for each λ_{max} that was resolved. The data treatment can also remove from the spectra differences due to nonuniformity in particle size. If the appropriate derivative segment is selected, a multicomponent absorption signal (fig. 4*B*) can be resolved into its individual absorption signals, with a minimum at the wavelength where the λ_{max} in log (1/R) (absorbance) spectra would have been had the band been resolved.

References

AOAC.
1980 Association of Official Analytical Chemists, Methods of analysis. 13th ed. Arlington, VA.

Barnes, R. F.
1973 Laboratory methods of evaluating feeding value of herbage. *In* G. W. Butler, and R. W. Bailey, eds., Chemistry and biochemistry of herbage. Academic Press Inc. Ltd. London.

Kaye, W.
1954 Near-infrared spectroscopy. I. Spectral indentification and analytical applications. Spectrochim. Acta. 6:257.

Kaye, W.
1955 Near-infrared spectroscopy. II. Instrumentation and techniques. [A review.] Spectrochim. Acta. 7:181.

Kowalski, B. R., ed.
1977 Chemometrics: theory and applications. ACS Symposium Series No. 52. American Chemical Society, Washington, DC.

Maddams, W. F.
1980 The scope and limitations of curve fitting (to vibrational band systems). J. Appl. Spec. 34:245.

Martens, H., and H. Russwurm, Jr., eds.
1983 Food research and data analysis. Applied Science Publishers. New York.

Mohsenin, N. N.
1984 Electromagnetic radiation on properties of foods and agricultural products. Gordon and Breach Science Publishers, Inc., New York.

Moore, J. E. and G. O. Mott.
1973 Structural inhibitors of quality in tropical grasses in anti-quality components of forages. *In* A. G. Matches, ed., Anti-quality components of forages. Crop Science Society of America, Madison, WI.

Norris, K. H., R. F. Barnes, J. E. Moore, and J. S. Shenk.
1976 Predicting forage quality by IR reflectance spectroscopy. J. Anim. Sci. 43:889.

Wheeler, J. L. and R. D. Mochrie.
1981 Forage evaluation: concepts and techniques. American Forage and Grassland Council, Lexington, KY, and CSIRA, East Melbourne, Victoria, Australia.

Willard, H. H., L. L. Merritt, and J. A. Dean.
1974 Instrumental methods of analysis, D. Van Nostrand Company, New York.

Animal Response Prediction

W. R. Windham and S. W. Coleman

Introduction

The ultimate criterion for assessing the quality of forages is their potential to support animal production, either production per animal or production per unit of land. There is a need for a rapid assessment of feed quality which will accurately predict animal performance. Today, NIRS offers the most potential as a rapid, nonconsumptive technique for estimating chemical and biological parameters of forage quality, including digestibility. However, high precision in predicting animal gain on forages by NIRS would not be expected because many factors are involved which are external to the chemical information in the feed sample and hence not detected by NIRS. One method to improve precision is to derive different equations for different feeds (that is, in terms of maturity, species, and so forth). An alternative method is to derive one general (universal) equation for all feed types. However, precision is likely to be lower with the second method and a bias may result from either method. A correction factor could be calculated for the bias source and be applied when each feed type is predicted from the general equation. An alternate method of removing bias can probably be achieved by spectral transformations using high speed computers.

The universal equation method has appeal to those predicting animal response since it is difficult if not impossible to calibrate for each type of feed, experiment, or other factor which may cause bias. The objectives of this report are to discuss factors that influence the accuracy and precision of NIRS predictions of digestibility.

Precision

Those accustomed to predicting chemical constituents such as protein content of grains and forages with NIRS expect a high R^2 and low standard error of prediction (SEP) for the equation. Although bioassays are typically low in precision an equation based on bioassays may be very good in estimating the actual attribute of the feed. In fact, accuracy may not be adversely affected at all, even though precision is lower than desired.

The precision of NIRS calibration based on bioassays is low for two primary reasons: (1) bioassays can vary according to environmental conditions and the inherent characteristics of the feed and (2) the factors influencing the bioassay results may or may not be chemicals

with specific compositions and spectral properties (for example, fiber and lignin, which influence digestibility).

Eckman et al. (*1983*) reported NIRS prediction errors for animal intake and digestible energy (DE) of 11 g/$W^{0.75}$/day and 0.2 kcal/g, respectively. NDF and protein were also used to predict intake and gave a lower standard error (7.7 g/$W^{0.75}$/day) than the NIRS method; however, NIRS was more precise than NDF alone. In vitro dry matter digestibility also gave lower SEP for estimating DE than did NIRS. These diets were fed to sheep and consisted of a mixture of concentrate, hay, and silage.

Accurate digestibility data may be obtained by minimizing all sources of variation in the analysis of forage quality, as discussed in appendix 3. Usually, the precision of predicting forage quality by NIRS is higher than the precision of in vitro and or in vivo digestibility analyses on which NIRS calibration equations are based. Therefore, predicted errors are more a function of the digestibility trial than the NIRS data. To obtain an adequate number of samples with digestibility values for calibraton and/or validation, one must use data from numerous trials. As such, the capability of NIRS to accurately predict digestibility will depend on the cause of the variation in the data within and between trials.

Because of the variation in bioassay data, one should not expect standard errors of calibration (SEC) and SEP as low as, and R^2 as high as, those associated with data obtained by other laboratory procedures (that is, data on CP NDF, ADF, and so forth). For only random errors, Shenk (unpublished) proposed the following equation to estimate the laboratory error for a given procedure:

$$R^2_{NIRS} = 1 - \left(\frac{2(SEL)}{SD}\right)^2$$

where:

R^2_{NIRS} = coefficient of determination by NIRS,
SEL = mean standard error of all samples in the in vitro procedure (standard error of laboratory), and
SD = standard deviation of the analytical or bioassay data in the sample set.

Expected R^2_{NIRS} values from different SELs and SDs of calibration and/or validation for in vitro sample sets are shown in table 2. With a small SD and high SEL, R^2 is insignificant. However, as the SD of the sample set increases and SEL decreases, a significant R^2 can be obtained. In addition, the above formula can be manipulated to estimate SEL from SD and R^2. SEC is generally about two times SEL.

Table 2.—*Expected NIRS coefficient of determination (R^2) from mean standard error of the in vitro analysis (SEL) and standard deviation of the calibration and/or validation sample set*

	Standard deviation				
SEL	3	5	7	9	11
1.0	0.56	0.84	0.92	0.95	0.97
1.5	.00	.64	.82	.89	.93
2.0	.00	.36	.67	.80	.87

Variation Within Trials

The variation of IVDMD values within a trial is generally small. However, failure to minimize all sources of random variation will increase the magnitude of the errors. The greatest sources of variation are the random analytical errors of weighing, sample transfer, and dry matter (DM) determinations. This is especially true if one is using the smaller sample size (250 mg). Correction of "as-is" analyses with accurate DM determinations will decrease the percentage analytical error of determining DM digestibility. When proper gravimetric techniques are used, the random analytical error will usually contribute little to the occurance of predicted significant *t* values with the NIRS equation.

Nonrandom variation (bias) also can occur within trials because of errors in determining the indigestible inoculum blank residue. If the inoculum is not properly agitated during inoculation of substrates, significant differences in the amount of inoculum blank residues will occur. Without proper agitation, the heaviest particular matter in the inoculum will settle to the bottom, so the amount of residue will continually increase as the inoculum is used. Marten and Barnes (*1979*) recommended that six inoculum blanks be interspersed throughout the forage samples and that the average DM residue be used in calculating IVDMD values. In our laboratories at the Russell Research Center in Athens, GA, and at the Livestock and Forage Research Laboratory in El Reno, OK, duplicate inoculum blanks are placed in each test tube rack (24 tubes), and the mean residue is used to calculate IVDMD values for that given rack. Regardless of how the blanks are interspersed, proper agitation of inoculum is essential to decrease nonrandom variation.

Variation Between Trials

The single greatest source of uncontrolled random variation between trials is the inoculum. The magnitude of random variability has been reported to be decreased by increasing the length of incubation period (Baumgardt and Oh *1964*), supplementing with nitrogen in the in vitro system (Alexander and McGowan *1966*), and adhering to strict standardization of the donor animals as well as the procedures for procurement and processing the forage samples. The importance of standardizing the treatment of the donor animals cannot be over emphasized. A portion of this variability may be removed by increasing the number of in vitro trials.

Standard forage substrates in duplicate or triplicate are also recommended for inclusion in each trial to verify in vitro results. In our laboratories, we use duplicate standard substrates in each test tube rack. The use of standards allows us to correct results from trial to trial and to determine when an entire trial should be discarded. However, no information exists on the effect of corrected IVDMD values on calibration and validation of NIRS.

Bias

Minson et al. (*1983*) reported systematic bias in using NIRS to predict animal intake, digestibility, and crude protein content of a group of forages of different species, maturities, and physical form (for example, long hay vs. pelleted hay). However, most of the bias they observed could have been due to particle size differences since log (1/R) was the only math treatment used.

NIRS cannot be expected to detect differences in feed intake and digestibility which are caused by physical factors, for example, grinding and pelleting effects. Minson et al. (*1983*) observed that pelleting reduced the amount of reflected energy, but one would not expect this reflective difference when the comparison was long versus coarsely chopped or even finely chopped material. The heating and pelleting process probably altered the structure of the forage, giving rise to particle size differences. Pelleting usually manifests itself in increased intake and reduced digestibility due to faster rate of passage caused by physical, not chemical, action. This interpretation is supported by the data of Minson et al. (*1983*), who found that pelleting reduced DMD by 4.2 percent, whereas NIRS predicted only a 1.3 percent reduction. However, we do not perceive a frequent need, especially in research, to discriminate between pelleted and chopped forages. Industry might have need for predicted values of feeds regardless of physical form, but generally an adjustment factor could be added. The problem with the 19-filter instrument that Minson et al. (*1983*) used is its inability to mathematically correct for particle size, drying regime, and environmental factors which influence the spectral baseline. With a spectrum obtained from a monochromator or scanning filter instrument equipped with adequate computational power, math treatments such as derivatives or Fourier transformation can be used to help correct systematic bias. Minson et al. (*1983*) also stated that NIRS "did not fully distinguish the higher level of organization of the cells walls in the mature forage." Perhaps the organization caused particle size differences which resulted in the bias. Almost all quality characteristics we estimate in forages are influenced by maturity; hence there is a need for a range of maturities of samples on which to develop a linear relationship during calibration.

Statistical models to remove the effects of particle size, species, and maturity, animal trial, season when fed, and so forth, are probably required when single wavelength filters are used. However, when the spectrum is available, more rational means of baseline correction (such as derivative math) become available.

Data Set Selection

Another major factor in successful NIRS analyses is the proper selection of calibration samples. The NIRS technique is strictly empirical; therefore, the accuracy of predicted results can never exceed that of the measurements with which the instrument is calibrated. Because the technique relies on multiple reflectance measurements to correct for interference from the different components in the sample, the calibration base samples are critically important. The question as to how many samples are required to accomplish calibration is hard to answer. Shenk et al. (*1976*) stated that "the simplest answer is to use as many samples as possible covering all variables that may affect the chemical and physical composition of the forage." The number of samples selected in the calibration base equation is dependent on the intended use of the equation, that is, whether it is to be used to predict a finite or infinite population. Shenk et al. (*1976*) proposed that in practical terms, 50 samples might be considered a minimum for calibration selection. Another rule of thumb which can be used for large populations is to select 10 to 20 percent of the total population.

Regardless of the percentage used, if the calibration sample set represents the population to be predicted and if the appropriate math treatments are used, little difference in the base calibration and validation data (bias) will be found. However, if the equation is to be used on a larger and highly diversified population, the higher percentage selection should yield more accurate predicted data and result in fewer significant *H* values.

Once a broad-base equation has been developed, samples can be added to the calibration set over time and the equation recalibrated for even wider use. Using a high percentage of the total sample selected during initial calibration should decrease the number of samples necessary during recalibration to accurately predict a larger and more diversified population.

The next factor which must be considered in calibration equation development is the method of sample selection. (See "Populations" (under "Calibration") for suggested approaches.)

Regardless of the number of samples and the method of sample selection, the calibration data must be accurate and must represent the population to be predicted. Errors in predicting animal response are generally greater than those in predicting chemical composition. This may be a result of error during collection of the animal response data used in calibration, and much of that error may be due to only a few outliers. With in vivo response data, it is usually impossible to retest suspect samples; therefore, one should check to determine if elimination of the outliers will improve the correlation. One may also eliminate outliers in IVDMD data to see if correlation is improved, but generally these outliers have high SELs and should be rerun. If one is predicting multiple constituents (for example, DM, CP, and IVDMD) and all have signficant *H* values for a given sample, that sample is likely to represent variation. In addition, NIRS prediction of digestibility depends on the assumption that all sources of random variation within and between trials have been kept at a minimum and that the same IVDMD method was used in calibration and validation of the population. Any bias in the original IVDMD data will lead to bias in predicted digestibility. The NIRS user should always remember that calibration is simply a mathematical technique used to teach the computer to accurately predict the quality components of forages based on the relationship between their NIRS spectra and chemical constituents. Therefore, one should not expect NIRS to accurately predict any sample or population of samples whose characteristics were not included in calibration.

References

Alexander, R. H., and M. McGowan.
1966 The routine determination of in vitro digestibility of organic matter in forages. An investigation of the problems associated with large scale operation. J. Br. Grassl. Soc. 21:140-147.

Baumgardt, B. R., and H. K. Oh.
1964 Evaluation of forages in the laboratory. IV. Within and among trial variability of the Wisconsin artifical rumen procedure. J. Dairy Sci. 47:263-266.

Eckman, D. D., J. S. Shenk, P. J. Wangsness, and M. O.
1983 Westerhaus. Prediction of sheep responses by near infrared reflectance spectroscopy. J. Dairy Sci. 66:1983-1987.

Marten, G. C., R. F. Barnes.
1979 Prediction of energy digestibility of forages with in vitro fermentation and fungal enzyme systems. *In* W. J. Pigden, C. C. Balch, and M. Graham, eds., Standardization of analytical methodology for feeds, p. 49-60. International Development Research Center, Ottawa, Canada.

Minson, D. J., and K. L. Butler, N. Grumitt, and D. P. Law.
1983 Bias when predicting crude protein, dry matter digestibility and voluntary intake of tropical grasses by near-infrared reflectance. Anim. Feed Sci. and Tech. 9:221-237.

Shenk, J. S., W. N. Mason, M. L. Risius, and others.
1976 Application of infrared reflectance analysis to feedstuff evaluation. *In* P. V. Fonnesheck, L E. Harris, and L. C. Kearl, eds., First International Symposium on Feed Composition, Animal Nutrient Requirements and Computerization of Diets. p. 242-247.

Validation of NIRS Results by Chemical Analysis

F. E. Barton II

Periodically, it will be necessary to validate an NIRS result in the laboratory. While on the surface this may seem simple, there are a number of factors which complicate the process. The obvious way to validate the NIRS result is to analyze the sample by the method that had been used to develop the calibration equation. In theory, this should validate or, for some reason, known or unknown, not validate the instrumental result. This means that either the sample does not fit the calibration or that the wet chemistry method does not yield the same result on this sample as it statistically should have for supposedly similar samples in the calibration set. The reasons are detailed in appendix 2. The factors that complicate the validation process are threefold. First, an empirical procedure will never yield a "right" answer; second, the NIR spectrum and chemical analysis were not run on the same sample; and, third, the resultant value, predicted based on the calibration samples, may not have the same relationship to a nutritive entity as the calibration samples or the ones you are analyzing with NIRS. These factors were pointed out, along with the statistical implications, by Lucas and Smart (*1961*) and Lucas et al. (*1961*). These are not new ideas; they have been recognized for over a century.

How then can we be certain, or at least maximize the possibility, that we can validate the result of an NIRS analysis in the laboratory? In theory, this only can be done by a procedure which measures the same nutritive entity in all samples the same way. Research is being conducted to find better analytical methods, but in the interim there are some things we can do to raise our confidence in the chemical results and in turn validate the NIRS result. First, we could use a different method. This is difficult with empirical procedures, but the analysis could be run by a different analyst or in a different laboratory. The point here is that no one runs a given procedure the same way. If there is an alternate procedure, such as dye binding in place of macro- or micro-Kjedahl for protein, that is better. In the case of fiber analysis, the Fibertec extraction unit could be used to check a conventional analysis.

Second, a certified standard method could be used. In this case the manual "Methods of Analysis" (AOAC *1980*) would be the reference of choice since all of its procedures have been checked for between-laboratory and within-laboratory errors. The certified procedure can be considered an alternate method as well. The use of a certified procedure implies that the analyst follow the procedure exactly.

Third, a qualitative or semiquantitative method could be used to evaluate the sample. These types of procedures are also discussed earlier. Here, the analyst will ascertain that the sample does fit or for some reason, does not fit the calibration. These methods will give information about the sample. Another aspect of acquiring information is to examine the other compositional values determined for the sample. These results will either be consistent or not. By these approaches the analyst will gain some insight as to the validity of the result.

Fourth, if all the samples in a set seem to be in question, an alternate set of samples that are known to fit the calibration can be used to validate the result. Many times the wrong equation file will have been used and even inadvertently renamed. These errors do occur, and the use of a second known file can eliminate them.

Finally, a reference set of samples can be used. These samples are ones that have been analyzed repeatedly by a conventional method so maintained that they undergo little to no change. Additionally, the laboratory chemistry data can be updated, assuming the supply of the sample is sufficiently large. In the research environment, keeping such a set or sets of samples is difficult. In the commercial or routine laboratory setting, it is mandatory to have a reference sample set. These samples are needed in addition to a ceramic or polystyrene standard. The reference set can be used as the alternate set above to validate an NIRS result. In addition, this set must be run routinely and a record maintained of the results so that condition of the instrument and the calibrations can be continually monitored.

References

Association of Official Analytical Chemists, AOAC.
1980 Methods of analysis. 13th ed., Arlington, VA.

Lucas, H. L. Jr., and W. W. G. Smart, Jr.
1961 Chemical composition and digestibility of forages. *In* S-45 Report. North Carolina State College, Raleigh.

Lucas, H. L., Jr., W. W. G. Smart, Jr., M. .A. Cipolloni,
1961 and H. D. Grass. Relations between digestibility and composition of feeds and foods. *In* S-45 Report. North Carolina State College, Raleigh.

Populations

S. M. Abrams

The population to be represented by the calibration set may be either finite or infinite. A finite population has defined boundaries which limit the population to a specific number, whereas an infinite population has no such boundaries. An example of a finite population would be all the plants in a specific forage breeding experiment. An example of an infinite population is all the alfalfa hays fed to livestock in the United States; there is no limit as to time or number of hays. An infinite population is generally more hetergeneous than a finite population, and will require a larger sample base for calibration. In addition, unlike finite populations, infinite populations may not always be adequately represented by the calibration set. For these reasons, the standard error of measurement for NIRS analysis is normally higher than is observed with finite populations.

In selecting samples from a population, two methods may be used: structured or random. Structured sampling is based on some a priori knowledge about the population. For example, when the population consists of forages harvested at three maturities, proportional sampling from each maturity ensures that maturities are equally represented in the calibration set, and that calibration will not be biased, as it might be if samples were randomly selected. Random selection is generally appropriate when no a priori information suggests the need for structured sampling.

A third method of selecting samples is on the basis of spectral characteristics. This is a form of structured sampling, whereby spectra are first obtained on all samples to be analyzed. The spectra are then grouped into clusters by spectral characteristics, and random samples from each cluster are used to establish the calibration set. Such a technique is scientifically appealing, because it suggests that the entire range of chemical variability, as expressed by NIR spectra, is expressed in the calibration set. However, preliminary research at University Park PA, and Beltsville, MD, indicates that such a technique offers no improvement over random sampling as a sampling method.

The essential point about calibration sample selection is that the samples chosen should represent the range of characteristics (chemical, physical, botanical, and so forth) present in the population of interest. Thus, a calibration equation developed from a sample set with a dry matter range of 90 to 95 percent should not be used to analyze a population with a dry matter range of 85 to 95 percent. Similarly, calibration equations should not be derived from legumes if the ultimate purpose is to measure the composition of grasses. This conforms to one of the cardinal rules of statistics, that is, do not extrapolate beyond the range of available information.

Equation Development

M. O. Westerhaus

Equation accuracy can be limited for three reasons. First is the accuracy of the reference method in analyzing the quality of the calibration samples. Random measurement errors will have minimal effect on the calibration, but any systematic errors will have a major effect. Second, the spectral information available may not be unique to the chemical bonds of interest. Chemical bonds absorb in several regions of the spectrum, the near infrared region showing the overtones and combination bands. Also, most regions of the spectrum show the effects of absorption by more than one chemical bond. Often, wavelengths corresponding to chemical bonds not of interest appear in equations to counteract the effect of other absorbers at the primary wavelengths. Furthermore, enough information must be present to separate spectral changes due to chemical changes from spectral changes related to changes in the physical properties of the sample, particularly particle size. Finally, even though the information is present, the calibration program may pass over the appropriate transformation of the NIRS data and optimal selection of wavelengths. The goal of the calibration procedure is to find the best fitting equation for the samples in the calibration set. A perfect fit of the samples, however, would not be desirable because the fit would apply not only to the samples but also to the errors. Yet, as more and more math treatments and wavelength combinations are evaluated, the calibration program is more and more likely to fit the errors along with the data.

Data transformations serve two primary functions: noise reduction and isolation of spectral information relating to sample chemistry. Noise reduction is usually accomplished by a simple, running-average smoothing of neighboring wavelengths. This procedure results in a small loss in spectral resolution and a large decrease in random noise. An alternative noise reduction method utilizes Fourier transforms to represent the spectra as sums of sines and cosines. The high frequency sines and cosines are assumed to represent random noise. If they are set to zero and the remaining sines and cosines combined, the resulting spectra will have less noise than the original while maintaining the original resolution. Since the noise on commercially available monochromators is very low, little smoothing is needed and a small running average is usually sufficient.

It is generally agreed that log (1/R) is the function of reflectance most linearly related to sample composition in the near infrared. However, the log (1/R) measurement at one wavelength is also responsive to the relative amount of the light scatter in the sample. Light scatter is related to the size and shape distribution of the ground sample particles (Birth *1976*). This is why sample handling and sample grinding must be standardized.

The primary effect of particle size on the spectra is a simple offset. A function that is unaffected by this offset is a linear combination of two or more wavelengths whose coefficients sum to zero. This function is precisely what derivatives, or finite differences, perform.

A better model of the effect of particle size is a combination of offset and a multiplicative factor. Once derivatives have been applied, the multiplicative factor remains. A function that is unaffected by this factor is a quotient (a derivatized data point divided by a reference-derivatized data point). The difficulty here is to find a reference wavelength that is sensitive to particle size and insensitive to sample chemistry.

Fourier transforms can be applied at this point to reduce the high intercorrelation normally found between adjacent NIRS data points. Calibration programs applied to Fourier coefficients usually result in calibration errors of the same size as errors resulting from calibrating without the Fourier transforms, but with more terms.

The principal-components method is another way of transforming spectra into a series of linear combinations of wavelengths. The first principal component is the linear combination of wavelengths that varies the most from sample to sample in the calibration file. The second principal component is the linear combination that is independent of the first component and varies the most from sample to sample. Thus, most of the variation in the calibration set can be expressed in a few terms for each sample. Both principal-component analysis and Fourier transforms divide the information at a specific absorption peak among several terms. This usually results in more terms being used in the calibration equation.

Two regression techniques can be used to estimate the equation coefficients and measure the goodness of fit. The standard method is least squares, which minimizes the sum of the squared residuals (SSR). The second method, ridge regression, minimizes the sum of the repeatability of the equation (measured on a file on repeated scans of one sample) plus the SSR. The effect of ridge regression is to discourage selecting wavelengths associated with high measurement noise or large equation coefficients. Once the wavelength selection is made, ridge regression will compute slightly smaller coefficients than least squares. Some shrinking

of the least squares coefficients is desirable, since least squares estimates are too large when the *x*-variables (log (1/R)) are measured with error. With the low noise of current scanning monochromators, however, ridge regression equations differ only slightly from least squares equations.

Wavelength selection is the most time consuming and potentially the most dangerous part of calibration. It is time consuming because thousands and sometimes millions of different wavelength combinations must be evaluated through a regression. The danger comes from selecting the best fit out of such a large number of choices. The best fitting wavelength combination selected from such a large number of choices usually fits the random errors as well as the model. Two methods are used to pick the wavelength combinations. The more straightforward approach is to try all possible combinations of two, three, or more wavelengths. However, the time required for trying more than three terms is prohibitive. The alternative to testing all possible combinations is the modified stepwise approach. The first term is simply the best fitting wavelength. The second term is one member of the best fitting pair whose other member is fixed as the first term. The second term is then fixed, and an attempt is made to find a term that fits better than the original first term. Each member of the pair is rejected in turn until no further improvements are found. These two terms are fixed as an attempt is made to find a good third term. Then each term is rejected one at a time as an attempt is made to find a better set of three terms. The same procedure is then extended to find sets of four, five, or more terms.

Reference

Birth, G. C.
1976 How light interacts with food. *In* Quality detection in foods, Amer. Soc. Agric. Eng.

Interpretation of Regression Statistics

M. O. Westerhaus

The preferred statistics to use for evaluating calibration regressions are based on validation samples separate from the calibration samples. The regression statistics, however, do provide additional information and may be the only statistics available when not enough samples are available for validation.

Before the regression statistics are evaluated, the residuals should be examined for samples with large *t* or *H* values. A large positive or negative *t* value indicates that the laboratory data could not be fit. Often, a reanalysis of the sample by the laboratory results in values that fit much better. Large laboratory errors should not happen often if multiple laboratory analyses are performed and checked for close agreement. A large *H* value indicates that the NIRS data used in the equation differed dramatically from the NIRS data of the other samples. Such a sample should be rescanned. If the two scans agree and the sample belongs in the population, then it should be retained as a very important sample. If the scans disagree, then the first one was probably a mistake and should be discarded. Samples with large *H* values have high leverage, and are usually fit very well. Thus, it is rare to find a calibration residual with both large *t* and *H* values.

The standard error of calibration (SEC) describes how well the calibration samples were fit. The lower the SEC, the better the fit. SEC will be zero in an equation that has one term for every sample. Such a low SEC is a clear case of overfitting, and means that even the random measurement errors are being fit. If the NIRS measurements and calibration process were error free, SEC would equal the laboratory repeatability error of the calibrated variable (SEL). In practice, however, NIRS data are measured with error and the calibration process is imperfect. These errors raise the smallest SEC attainable without overfitting. The rule used at University Park, PA, is to select an equation with an SEC about twice the SEL.

The coefficient of determination (R^2) is the proportion of variability explained by the regression equation. It varies inversely with SEC. An R^2 of 1.0 corresponds to an unrealistic SEC of zero. R^2 is approximately equal to 1.0 minus SEC squared divided by the standard deviation of the data squared. Low R^2 values are often an indication that the laboratory data are imprecise. If the SEL were one-fourth the size of the standard deviation, the University Park rule would suggest selecting an equation with an R^2 of 0.75.

The final regression statistic to be considered is the *F* statistic. High *F* values indicate that the associated coefficient is significantly different from zero. Small *F* values indicate that the coefficient might contribute little to the equation except to fit the random errors. The probability that the observed *F* value was obtained solely through chance does not follow the standard *F* tables since the observed *F* is selected as the maximum out of all the wavelength combinations considered. As the number of choices increases, larger *F* values are needed to signify a coefficient fitting more than just random errors. The rule followed at University Park is to reject equations with *F* values less than 10. Note should be made that as the number of samples increases, NIRS variables making a small contribution to the overall fit of the equation will have large *F* values. Although equations with fewer terms may have slightly lower accuracy, they often have better repeatability, are easier to transfer, and are less likely to be overfit. Although only one equation can be the best, several other equations usually exist that can perform satisfactorily.

Finally, the wavelengths should be considered in the equation selection process. For one- or two-term equations, the wavelengths should correspond to known absorption peaks related to the laboratory analysis. For multiterm equations, the wavelengths are so interdependent that interpretation of individual wavelengths is difficult. It is, however, useful to consider whether wavelengths in the water regions are part of an equation.

Validation

M. O. Westerhaus

There are three steps in validating an equation. Validation may be performed on a subset of the calibration set reserved for that purpose to aid in the equation selection process. The selected equation should subsequently be validated with unknown samples during routine NIRS analysis.

The first step in using validation to select equations is to examine residuals for large *t* and *H* outliers. A large *t* value for a few samples would indicate that the laboratory values did not represent the samples at the time the scan was taken. These sample measurements should be omitted from the analysis and replaced with new measurements of the same samples. A few bad sample measurements can greatly influence the validation statistics. Many validation samples with large *t* values would indicate that overfitting has occurred and that the equation is specific to those samples in the calibration set.

A large *H* value for a few samples indicates that their spectra are different from the spectra of the calibration samples. These samples should be examined for possible omission only if their *t* values are also large. Many large *H* values indicate overfitting.

The next step is to use the validation statistics. The standard error of prediction (SEP) is a true indication of the performance of the equation on unknowns from the same population. The rule followed at University Park is to select the equation with the smallest SEP. Unlike SEC, which must decrease with each additional term, SEP only decreases until overfitting becomes important and forces it to increase. Another useful statistic is the slope of the regression line relating the determinations to the laboratory values. Slopes should be near 1.0. If markedly different from 1.0, they indicate the high and low values will be consistently under or over estimated.

Once an equation is selected, it should be refit using the same wavelengths and math treatments with all the samples. The final step is to validate the equation with unknown samples during routine analyses. The most likely problem encountered would be a shift in the sample population from that used in the calibration. This shift can be detected by monitoring the *H* statistic on the unknown samples. If the average *H* becomes greater than 3.0, the population has probably shifted so much that recalibration is necessary.

Recalibration
M. O. Westerhaus

Recalibration is necessary whenever the instrument or the sample population changes. Samples representing the current population are selected, analyzed in the laboratory, and added to the original calibration samples. Problem samples, such as moldy or contaminated samples, should not be added. The entire calibration procedure is then repeated.

The preferred method of selecting samples for recalibration is to first obtain laboratory values of all samples during routine NIR analyses. Then samples with large differences between laboratory values and NIRS values should be reanalyzed and, if the differences are still large, reserved for recalibration.

If the differences do not average zero, a bias exists. If the bias is the major component of the differences, recalibration can be simplified by fitting the original wavelengths and math treatments to the augmented calibration file.

Calibration Transfer
J. S. Shenk

Calibrations are sensitive to the small wavelength and photometric differences normally found among instruments. The wavelength effect can be caused by slight differences in grating alignment or, in the case of filter instruments, in the center wavelength tolerance. Refractive index is an additional source of variation in tilting filter instruments. Sources of photometric differences among instruments are stray light, specular component, nonlinearity in detector response and signal amplification, and small but important differences in the overall optical system.

These small but important differences, present even among instruments from the same manufacturer, prevent the direct transfer of a calibration from one instrument to another. The limitation this imposes on the practical use of NIRS is obvious. Calibrations are expensive to develop and should be used on many instruments to be cost effective.

The software developed by the Network to enable calibration transfer has been briefly described earlier. Wavelength alignment between monochromators can be corrected by using the polystyrene test in the program SCAN. A set of transfer samples is utilized by the program TRNSFR to make even finer adjustments in wavelengths and to correct for photometric differences among instruments.

Transfer samples are used to derive the linear corrections that make the transformed data of one NIRS instrument more like those of another instrument. To be most effective, the range of transformed NIRS data in the transfer samples must cover the range of transformed NIRS data in the calibration samples at the wavelengths used in the equation. The range of transformed data in the transfer samples can be wider than that of the calibration samples, provided that the linear corrections for the calibration samples are still appropriate for the transfer samples.

From 10 to 20 transfer samples can be selected from the calibration set, and they should be high and low in moisture and any other variables of interest. If calibrations for several products (for example, hay and haylage) are to be transferred, the transfer samples from each calibration set can be combined into a common transfer set. Samples outside the calibration set should be added if the range of moisture or other variables of interest is small.

It is important to prevent any changes in the transfer samples since all changes in the spectra are attributed to the instrument. There are two ways to maintain the integrity of a transfer sample. One way is to place the sample in a holder and seal it to minimize sample change over time. The sealing process prevents changes in the moisture level of the sample. Since NIRS is extremely sensitive to moisture, even a minute change in moisture will cause problems in transfer. Care must be taken to minimize the development of microscopic shifts or cracks in the sample surface due to settling of the sample over time, to prevent changes in sample chemistry due to sunlight exposure, and to prevent any type of fermentation reaction.

The other way to prevent changes is to seal the sample in an aluminum pouch, as described in "Sample Preparation," rather than in a sampler holder. Before a sample is used for transfer, spectra should be collected on four subsamples. Averaging the spectra of these subsamples will minimize transfer error between instruments.

Even when the best efforts are made to control sample errors, the samples will still change somewhat over time. The best precaution is to systematically collect the spectra for the calibrated instrument on a regular basis so that the spectra are always in the best condition for transfer.

Our research has shown that the equation with the lowest standard error of prediction on the calibrated instrument may not be the best equation to transfer to the uncalibrated instrument (Shenk et al. *1985*). For this reason we recommend that a series of equations be developed for each variable to be transferred. This multiple set of equations should be transferred and tested on the uncalibrated instrument. At least four equations with different math treatments and wavelength combinations for each variable should be tested.

Transfer of the equation can be accomplished by one of four techniques.

1. Adjust the analysis values on the uncalibrated instrument for slope and bias.

2. Adjust every term in the equation from the calibrated instrument for wavelength shift in the uncalibrated instrument, and recalculate the slope and intercept of each term for differences in NIRS data between instruments.

3. Recalculate the calibration variance-covariance matrix of the uncalibrated instrument to fit the determinations of sample quality of the calibrated instrument.

4. Transfer the spectra from the calibrated instrument to the uncalibrated instrument, add spectra and chemical data from new samples, and recalibrate the new instrument.

Each of these transfer methods has advantages and disadvantages.

Method 1 is simple and easy to accomplish if an appropriate set of samples is present to make the adjustment.

Method 2 is accomplished by the program TRNSFR and is our method of choice because it allows for changes in wavelength as well as adjustment in slope and bias of the NIRS data. This method of course will not work on a fixed filter instrument but is mandatory on a tilting filter instrument. The EVALU8 program is used to match determinations of sample quality between the calibrated and uncalibrated instrument.

Method 3 is not available at present but has good potential. Adjustment of the variance-covariance matrix may be a good feature; but if the transfer set is not properly chosen, this method may distort this delicate mathematical balance.

Method 4 may be conducted with the NIRTRN program and is sometimes useful after instrument repair.

Regardless of the method used for transfer, evaluation of the transfer is an essential part of the transfer process. A successful transfer is one in which the standard error of a difference between instruments is at a minimum without sacrificing accuracy of the analysis. The statistic calculated by the program EVALU8 to estimate both of these criteria is the lowest average SEC and SED. In general the SED of the selected equation should be as small as the standard error of the laboratory (SEL) for this analysis.

The evaluation samples can be chosen in one of two ways: (1) use the transfer set or (2) use another set of independently chosen samples. The disadvantage of using the transfer samples is that no estimation of the bias can be obtained. The advantage is that if the transfer samples were chosen properly, they are probably the best possible set of samples to make the evaluation. Using a second set of samples may appear to make the test more fair, but care must be taken that they are an independent subset of the calibration or the results could be misleading.

Transfer of calibration between two commercially available monochromators is no different from the transfer of calibrations between instruments of the same manufacturer. The transfer of calibrations between instruments of different types is more difficult. An example would be the transfer of a calibration developed with a PSCO 6350 to a 51A. The first consideration would be to restrict the monochromator in the calibration program to consider only the NIRS data of the 6350 that is common to the 51A. This means restricting the 6350 to the six filter segments available for the 51A. Also, it would be advisable to omit the first and last five data points from the calibration to allow for differences in filter characteristics. After the simulated calibration is accomplished, the remainder of the transfer process is the same as discussed above.

Successful transfers have been made between models 6350 and the 51A, but the number of terms that will transfer satisfactorily is usually less than the number that can be supported on model 6350. For example, if a protein equation is developed for a hay file containing eight or nine wavelengths, only three to six of these wavelengths can be transferred successfully. This phenomenon is due to differences in photometric response and bandpass between instrument types. Transfers from model 6350 to model 4250 have only been useful with one term equations. Transfers to fixed filter instruments have not been attempted, but they will be limited to log (1/R) math.

The best transfer is accomplished between instruments of the same model: 51A to 51A, 4250 to 4250, and so forth. Under this restriction, the calibration program must be set up to select terms only within filter segments, not across segments. Again, the elimination of NIRS data points at either end of the segment is advisable. The major limitation to transfer of calibrations among filter instruments is the alignment and coverage of the filters. Transfer of multiterm equations will require very close agreement between instruments and very close attention to the transfer procedures discussed above. If a successful transfer cannot be made, a step by step approach to consider all potential problems

Table 3—*Accuracy of transferring a multiterm equation developed on a master instrument to other instruments of the same model, with accuracy expressed as standard error of a difference between instruments*

Model/ Instrument No.	Equation terms[a]									
	CP[b]	ADF[c]	NDF[d]	Lignin	IVDMD[e]	P	Ca	K	Mg	DM[f]
Pacific Scientific 6350										
1	0.15	0.45	1.04	0.22	0.45					
2	.17	.45	.74	.25	.50					
3	.27	.54	.89	.30	.53					
4	.21	.38	.60	.23	.48					
5	.25	.43	.78	.28	.50					
Technicon 500	.22	.37	.71	.30	.55					
Pacific Scientific 51A										
1	0.32	.040	0.85	.035	0.66					
2	.27	.23	.92	.44	.47					
3	.25	.37	.99	.23	.54					
Pacific Scientific 4250										
1	0.15	0.69				0.001	0.008	0.005	0.002	0.20
2	.13	.73				.002	.005	.006	.001	.15
3	.18	.59				.001	.009	.004	.002	.23
4	.11	.72				.001	.003	.008	.001	.11

[a] Blanks indicate absence of terms in equation.
[b] CP = crude protein.
[c] ADF = acid detergent fiber.
[d] NDF = neutral detergent fiber.
[e] IVDMD = in vito dry matter digestibility.
[f] DM = dry matter.

should be used. Failure to transfer calibrations after eliminating the above problems is probably due to poor alignment of the filters.

Although there are many potential problems with the transfer process, these procedures make it possible for commercial companies to provide laboratories with factory calibrated instruments. Only when precalibrated instruments are made available will this computerized technology begin to reach its potential.

Table 3 demonstrates the accuracy of the transfer procedure. Data are not presented for transfer between unlike models because of the difficulties discussed above. Transfers were made using multiterm equations containing five to nine terms per equation. The calibrations were developed from hundreds of samples and transfers made with 15 to 30 samples. In practically every case, the SED between instruments was less than or equal to the normal SEL of the chemical analysis within a laboratory. This means that by using the transfer process, errors between laboratories can be reduced to the level of those within laboratories.

Reference

Shenk, J. S., M. O. Westerhaus, and W. C. Templeton.
1985 Calibration transfer between near infrared reflectance spectrophotometers. Crop Sci. 25:159-161.

Current Applications of NIRS Technology in Forage Research

G. C. Marten

NIRS has been successfully used by forage researchers engaged in experimental plant breeding, crop management, plant physiology, and ruminant nutrition. In all cases, successful use of NIRS has been accompanied by strict adherence to the four imperatives outlined by Shenk et al. (*1979*), which have become cardinal rules for transfer of this technology:

1. Make sure that the calibration samples adequately represent the unknown samples in the population to be analyzed;

2. Conduct accurate laboratory analyses (chemical or bioassay) of the calibration samples;

3. Select the correct data processing techniques to extract the maximum information from the NIRS spectra; and

4. Select the correct wavelengths (from two to nine, depending on the assay and the population to be measured).

In the section on the history of NIRS, Clark reviewed much of the literature concerning the successful use of NIRS by researchers. Therefore, only examples of successful transfer of NIRS technology will be given herein to illustrate some of the precautions needed and the magnitude of errors that one can expect during routine NIRS assays of forage constituents.

Forage Plant Breeding

Marum et al. (*1979*) reported the first use of NIRS to measure forage quality (six constituents) in a genetics and breeding program. They determined heritability of cell wall constituents and associated quality components in reed canarygrass (*Phalaris arundinacea* L.). They used from 50 to 72 samples chosen at random from the entire population of samples to be assayed during development of calibration equations with a scanning monochromator instrument. Correlations between laboratory and NIRS estimates were $r = 0.90$ or greater for two fiber fractions and crude protein and were $r = 0.73$ or greater for lignin, silica, and in vitro dry matter digestibility (IVDMD). Standard errors of calibration were usually less than 1 percent for all constituents except IVDMD (1.7 percent); these errors were lower than those commonly reported for the conventional laboratory assays.

Marum et al. (*1979*) found, however, that when they used calibration equations based on samples of one year to predict the forage quality of samples of the following year, the standard errors of prediction more than doubled. This finding substantiated the need to develop separate calibration equations for each year. Major changes in spectral properties of the reed canarygrass between years were attributed to differences in color of the samples between years (light brown vs. bright green), and these differences were traced to different methods of sample drying between years.

These results show that methods for processing, storing, and handling forage samples intended for NIRS analysis must be uniform. Whenever a specific sample preparation procedure is omitted from the calibration set of samples, either the calibration equation will need updating to accommodate the change or a completely new calibration equation will be needed to represent the new population. Further, if plants that lead to an unknown set of samples are grown in an environment not represented by plants grown for the calibration set, equations will usually improve by updating to accommodate the new environment. This has been the case in Minnesota's breeding or management programs for forage quality traits in perennial grasses, perennial legumes, legume-grass mixtures, and small grain crop forages. Maize stover is the only product for which yearly updating has not improved initially excellent calibration equations. Perhaps the reason for this is that maize stover is not as susceptible to environmentally induced changes as are forages that are not physiologically mature.

Updating can be easily accomplished by conventional laboratory assay of 30 or fewer samples from the new population. After spectral analysis, about 15 or fewer of these samples should be combined with those from the original 50 to 70 to provide the updated calibration equation. Then, the remaining 15 or fewer can be used to validate the updated equation.

Because the typical plant breeder seeks relative rather than absolute forage quality values in striving to improve specific traits within a population to be screened, bias concerns are kept to a minimum. The breeder is most concerned about the precise measurement of thousands of plants. Usually, only small amounts of forage are required by NIRS, and this requirement facilitates quality testing of single plants or even portions of single plants. The breeder will also be interested in knowing which plants are outliers to the sample population. Fortunately, available NIRS computer programs will identify outliers by "starring" unusual unknown samples. The breeder can then deter-

mine by conventional assay whether the starred sample has been inaccurately assessed by NIRS (indicating the need for its addition to the calibration set) and/or whether it has a truly unique forage quality. Because NIRS can easily provide the fast, precise, and outlier information sought by the plant breeder, plant breeding is the uppermost area of research for which this new technology can be recommended without need for additional development.

Forage Crop Management

Among studies showing the successful use of NIRS to analyze forage quality in cutting-management experiments are those involving four small-grain crops, alfalfa (*Medicago sativa* L.), and four perennial legumes analyzed simultaneously (Marten et al. *1983*b, *1984*). Table 4 compares the standard errors in those studies with the errors associated with conventional and NIRS analyses of 30 cool- and warm-season forages at the six laboratories of the NIRS Forage Quality Research Project Network (Templeton et al. *1983*).

Once adequate calibration equations are developed, NIRS can be used to rapidly analyze forage quality in forage research laboratories. For example, the five components in the Minnesota studies (Marten et al. *1983*a, *1983*b, *1984*) could be routinely analyzed 25 times faster by NIRS than by the conventional methods.

Because NIRS permits the analysis of about 30 or more samples per hour, forage quality can be monitored with small daily samples during growth of forages in management experiments; this capability allows treatment effects to be measured before harvest of entire plots.

Equations for predicting forage quality by NIRS must account not only for the methods used in sample processing, storing, and handling, as discussed earlier, but also for the growth conditions (which can vary with time and location) and harvest methodology. For example, if the calibration equations are initially developed with samples harvested by either sickle-bar or flail machines, they will be improved by updating with samples obtained by both types of machines if the two types are to be routinely used to gather samples (unpublished data, G. C. Marten). Usually, equations will be improved also by updating to represent new environmental (year and location) effects. Fineness of grind, type of grinder, and sharpness of blades in grinders also will influence NIRS spectral readings of forage samples. These variables should be kept as uniform as possible to avoid measurement errors. If forage quality of plant parts is to be included in a management experiment, each plant part of each species at each maturation stage must be adequately represented in the calibration sample set. If this precaution is not taken, greatly biased forage quality measurements will result (unpublished data, G. C. Marten et al.).

Table 4--*Standard errors and coefficients of determination (R^2) associated with NIRS quality analyses of forages grown in management experiments*

Forage/(Source)	Method of analysis	Standard error (%)					R^2				
		IVDMD[a]	CP[b]	NDF[c]	ADF[d]	ADL[e]	IVDMD	CP	NDF	ADF	ADL
Mixed (Templeton et al. *1983*)	Conventional	3.67	0.43	1.70	1.40	0.90					
	NIRS	.83	.56	1.12	.60	.63					
Small grain crops (Marten et al. *1983b*)	NIRS	≥3.11	≥.68	≥1.79	≥.98	≥.31	0.92	0.99	0.94	0.98	0.96
Alfalfa (Marten et al. *1984*)	NIRS	1.56	.42	1.46	1.30		.94	.93	.96	.97	
Four legume species (Marten et al. *1984*)	NIRS	2.06	1.00	2.23	1.70	.63	.97	.99	.98	.99	.98

[a] IVDMD = in vitro dry matter digestibility.
[b] CP = crude protein.
[c] NDF = neutral detergent fiber.
[d] ADF = acid detergent fiber.
[e] ADL = acid detergent lignin.

Forage Crop Physiology

Insofar as NIRS should be capable of measuring any organic constituent that constitutes about 1 percent or more of the dry matter in plants, it should be useful to forage plant physiologists as well as breeders and management researchers. The precautions outlined earlier to ensure calibration equations that adequately represent the population of plant material to be assayed and that represent uniform harvest, handling, and preparation procedures of samples also apply to measurement of physiologically important constituents.

Marten et al. (*1983*a) and Brink and Marten (*1985*) documented that NIRS was well suited for analyzing total nonstructural carbohydrates (TNC) in alfalfa roots. They reported coefficients of determination for TNC of $R^2 = 0.90$ to 0.99 and standard errors of NIRS analysis of TNC as low as 1.17. The comparable standard error of analysis of TNC by a cumbersome conventional procedure was 1.14.

Ruminant Nutrition

Norris et al. (*1976*) were the first to report that NIRS (scanning monochromator instrument) could reasonably predict forage quality in vivo. They found R^2 values of 0.78, 0.64, and 0.72, respectively, for dry matter digestibility, dry matter intake, and digestible energy intake of grass and legume forages by sheep, even though some of the in vivo animal response measurements may have been erroneous. In feeding experiments with sheep, Shenk et al. (*1982*) found that NIRS with a scanning monochromator predicted digestibility and intake of 60 field-cured temperate hays as accurately as did fiber and in vitro digestibility assays. Also, Eckman et al. (*1983*) concluded that a scanning monochromator NIRS instrument showed potential for predicting ruminant animal responses (digestible energy, dry matter intake, and digestible energy intake) to pure and mixed-forage-based diets as accurately as did conventional laboratory analyses.

Burdick et al. (*1981*) used a scanning filter instrument for NIRS prediction of in vivo digestible dry matter in bermudagrass hays after primary and secondary wavelengths had been initially selected with a scanning monochromator instrument. Standard errors of 1.78 to 2.54 were less than those obtained in animal feeding trials (3.1).

Less success in use of NIRS to predict forage quality in vivo has occurred with fixed filter instruments. Winch and Major (*1981*) found that the standard errors of predicting in vivo digestibility of grasses and legumes with a fixed filter instrument were usually unacceptable. Minson et al. (*1983*) reported substantial biases for NIRS prediction of dry matter digestibility or voluntary intake of tropical grasses associated with species, plant part, or physical form variables. However, they confined their measurements to 19 preselected wavelengths in a fixed filter instrument. Thus, it appears that the greater flexibility provided by scanning filter and especially scanning monochromator instruments is necessary for satisfactory prediction of animal response to forage-based rations.

Ward et al. (*1982*) used NIRS (scanning monochromator) to estimate the forage intake of animals grazing diverse arid or semiarid rangeland by analyzing esophageal fistula samples. They found that NIRS analyses of forb and browse species were as applicable as conventional chemical analyses for predicting intake.

Animal response data for the NIRS calibration equation must be obtained as accurately as data on chemical constituents if the equation is to adequately predict in vivo values for unknown samples. If accurate base data are used, scanning monochromator instruments can predict digestibility and intake of forages just as well as or better than in vitro procedures and in much less time.

Because NIRS will allow daily monitoring of forage quality in animal nutrition studies, diets can be quickly altered during large-scale feeding experiments to compensate for forage quality changes and thus keep constant the nutrient levels defined at the outset for specific treatments. Also, alterations in diets as they pass through the digestive tract can be assessed via fistula sampling, and the composition of fecal as well as feed samples can be easily monitored via NIRS.

References

Brink, G. E., and G. C. Marten.
1985 Analysis of alfalfa root carbohydrate concentration by near infrared reflectance spectroscopy. Crop Sci. [In press.]

Burdick, D., F. E. Barton II, and B. D. Nelson.
1981 Prediction of bermudagrass composition and digestibility with a near-infrared multiple filter spectrophotometer. Agron. J. 73:399-403.

Eckman, D. D., J S. Shenk, P. J. Wangsness, and M. O.
1983 Westerhaus. Prediction of sheep responses by near infrared reflectance spectroscopy. J. Dairy Sci. 66:1983-1987.

Marten, G. C., G. E. Brink, D. R. Buxton, and others.
1984 Near infrared reflectance spectroscopy analysis of forage quality in four legume species. Crop Sci. 24:1179-1182.

Marten, G. C., G. E. Brink, J. L. Halgerson, and D. R.
1983a Buxton. Near infrared reflectance spectroscopy prediction of forage quality and total nonstructural carbohydrates in legumes. Agron. Abstr., p. 135.

Marten, G. C., J. L. Halgerson, and J. H. Cherney.
1983b Quality prediction of small grain forages by near infrared reflectance spectroscopy. Crop Sci. 23:94-96.

Marum, P., A. W. Hovin, G. C. Marten, and J. S. Shenk.
1979 Genetic variability for cell wall constituents and associated quality traits in reed canarygrass. Crop Sci. 19:355-360.

Minson, D. J., K. L. Butler, N. Grummitt, and D. P. Law.
1983 Bias when predicting crude protein, dry matter digestibility and voluntary intake of tropical grasses by near infrared reflectance. Anim. Feed Sci. Tec. 9:221-237.

Norris, K. H., R. F. Barnes, J. E. Moore, and J. S. Shenk.
1976 Predicting forage quality by infrared reflectance spectroscopy. J. Anim. Sci. 43:889-897.

Shenk, J. S., S. M. Abrams, P. J. Wangsness, and H. W.
1982 Harpster. Predicting the feeding value of hay from chemical analyses and near infrared spectroscopy. Agron. Abst., p. 153.

Shenk, J. S., M. O. Westerhaus, and M. R. Hoover.
1979 Analysis of forages by infrared reflectance. J. Dairy Sci. 62:807-812.

Templeton, W. C., Jr., J. S. Shenk, K. H. Norris, and others.
1983 Forage analysis with near-infrared reflectance spectroscopy—status and outline of national research project. *In* Proc. 14th Int. Grassl. Cong., p. 528-531. Lexington, KY. Westview Press, Boulder, CO.

Ward, R. G., G. S. Smith, J. D. Wallace, and others.
1982 Estimates of intake and quality of grazed range forage and near infrared reflectance spectroscopy. J. Anim. Sci. 54:399-402.

Winch, J E., and H. Major.
1981 Predicting nitrogen and digestibility of forages using near infrared reflectance photometry. Can. J. Plant Sci. 61:45-51.

Extension Applications in NIRS Technology Transfer

N. P. Martin and J. G. Linn

Introduction

The range of applications of NIRS in extension programs is enormous. NIRS analysis not only provides a means for high-volume forage testing with a minimum of labor but also allows forages to be tested directly on farms or in other field locations. Combining this technology with computer ration balancing or other forage/nutritional programs in the field expands the extension educator's capacity for providing field training and education. However, its sophisticated technological requirements, high investment costs, and personnel support requirements may limit its usefulness. This section will present what can be done within reasonable inputs by agricultural extension.

Capabilities and Compatibilities in Extension Programs

The accuracy of NIRS technology depends on the degree of commitment made to each of the following instrument calibration rules:

1. Select a calibration sample set that represents the range of characteristics in the unknown population to be analyzed.

2. Use the very best chemical analyses and techniques possible on calibration samples.

3. Select the best wavelengths and best mathematical treatment of NIR spectra for equation calibration.

4. Use a technically competent operator.

Calibration Equations

The major contribution of NIRS to the livestock industry is its nonconsumptive, nonchemical, rapid-testing capability. Of importance to the livestock industry is the ability to test forages frequently and incorporate test results in ration formulations. To meet these objectives, NIRS technology cannot tolerate low accuracy; and it will be low if calibration is not understood and not conducted properly. Calibration is explained in great detail elsewhere in this handbook. However, educators also must recognize that calibration adequacy depends on selection of adequate calibration samples; proper drying, grinding, and chemical analysis of calibration samples; and adequacy of the NIRS in-

strument, microcomputer, and software. All steps used in preparing samples for calibration must be exactly repeated in the analysis of unknown samples. *Therefore, it is of paramount importance to define one's analytic objectives before developing or purchasing calibration equations.*

Scientists in Pennsylvania, Wisconsin, and Minnesota have successfully developed accurate calibration equations for testing the quality of cool-season perennial forage crops harvested as hay or hay-crop silage, of corn silage, high-moisture corn grain, and small-grain crops harvested as forages. These forages are successfuly being analyzed by NIRS for CP, ADF, ADF-N and NDF (the last tested in only Minnesota and Wisconsin). In addition, mobile NIRS units in all three States are attempting to analyze all forage except high moisture corn for calcium, phosphorus, potassium, and magnesium. Minnesota has also attempted to analyze hay for sulfur, but NIRS accuracy for this constituent is worse than that for other minerals. The Pennsylvania State Forage Testing Laboratory does not use NIRS for mineral testing because the accuracy of these calibration equations is less than that of conventional mineral analysis methods. Because NIRS functions by measuring H bonding with C, O, and N in *organic* constituents, any successful mineral analysis can be explained only by the association of a mineral concentration with an organic substance.

Initial research with NIRS for use in forage quality analysis indicated that the best calibration equations were derived when the population was precisely specified (for example, species specific). However, in developing equations for farm and Extension use, it is best that they be robust (that is, that they be developed with samples prepared according to any of a number of methods that might be used by individuals requesting forage analysis). For example, the equations used by the Minnesota mobile NIRS unit were developed with forage samples processed by two methods of drying (oven and microwave) and by two or three different cyclone mills. In addition, all three States have developed equations that include the harvest and storage techniques used for all species grown in the region. Not all farmers can accurately differentiate the species in a mixture; thus, hay and hay crop silage equations were developed with mixtures as well as with pure species expected to be analyzed. The statistics of each analysis for the Minnesota mobile unit are shown in table 5.

Minnesota's mobile NIRS unit, as well as Pennsylvania's and Wisconsin's, uses a scanning filter instrument (Pacific Scientific 51A). Thus, its accuracy cannot be expected to be as high as that of a scanning monochromator; however, standard errors of analysis compared favorably with those reported by various laboratories. Templeton et al. (*1983*) reported that the standard errors of NIRS analysis with scanning monochromators in the National Network were ±0.56, ±0.60, and ±1.12 percent for CP, ADF, and NDF, respectively. The standard errors of analysis in the Minnesota mobile unit were larger (table 6), but those for CP and ADF were usually less than the conventional chemical analysis errors between laboratories reported by Templeton et al. (±0.75 and ±2.24 for CP and ADF, respectively). The Minnesota unit's standard errors for NDF were usually larger, but nevertheless fairly close to, those reported for conventional chemistry by Templeton et al. (±1.82). Therefore, the NDF should be obtainable with a reasonable degree of confidence by NIRS. When chemical values were correlated to NIRS values, greater than 75 percent of the variation was explained with NIRS analysis except for ADF-CP in haylage and CP in corn silage. However, ADF-CP is not as accurately measured chemically as most of the other assays, and NIRS measurement cannot be accurate if the calibration samples have not been accurately assayed by conventional means.

Equation and Spectra Transfer

NIRS equations have been successfully transferred between instruments. Transfer between instruments reduces initial calibration costs. Shenk and Westerhaus (personal communication) were successful in transferring spectra files from Pennsylvania and Minnesota to their scanning monochrometer instrument and combining those two files with their own, multi-State, file (table 7). The standard error of calibration (SEC) for CP was higher for the combined equation (file 4) than for the separate equations. The Minnesota mobile unit's hay equations have also been transferred to another compatible scanning filter instrument. Thus, transfer of calibration information (both equations and spectra) among compatible instruments is possible; this capability reduced calibration costs, calibration time, and calibration differences due to variations among laboratories. Although potential exists for sharing NIRS calibration equations and spectra, transfer practices need more evaluation as do among-laboratory variations in conventional chemical analyses.

Equation transfer has great potential, but it applies only to a system in which one master instrument supports other slave instruments. The transferring of NIRS equations will eliminate the laboratory-to-laboratory variations of conventional chemical analyses now occurring and which were shown by Templeton et al.

(*1983*) to be greater than NIRS variations between laboratories. However, this scheme prevents operators of slave instruments from updating their equations to accommodate local needs. Updating must be done by the master instrument because it contains the base NIRS specta. At present, additional limitation to equation transfer from master to slave instruments is that the transfer can be made only between instruments manufactured alike. Thus, NIRS equipment currently must be of the same brand and have spectral properties as similar as possible.

Table 5.—*Information used in equation development of forage quality by the Minnesota mobile NIRS unit*

Forage quality assay	Correlation data					Wavelength[e] No.
	N	R^{2a}	SEC[b]	SD[c]	Math[d]	
Hay						
DM[f]	115	0.87	0.66	2.55	1	5
CP[g]	102	.98	.63	4.05	1	7
ADF[h]	102	.91	1.43	4.85	1	7
ADF-CP	59	.71	.36	.67	1	5
NDF[i]	101	.93	2.55	9.42	2	7
Hay crop silage						
DM	76	.85	.97	2.55	1	3
CP	51	.90	.85	2.64	1	4
ADF	51	.84	1.60	4.04	1	6
ADF-CP	55	.59	.81	1.26	1	6
NDF	48	.62	3.40	5.55	1	6
Corn silage						
DM	76	.92	.60	2.16	2	3
CP	61	.91	.47	1.52	1	5
ADF	61	.94	1.92	7.53	1	3
NDF	60	.93	2.22	8.65	1	3
High moisture corn						
DM	156	.96	.37	1.92	2	3
CP	137	.89	.32	.97	1	5
ADF	137	.98	.49	3.14	1	6
NDF	133	.97	.94	5.45	1	6

[a] Squared coefficient of multiple determination from the least squares regression of known forage quality values on NIRS values.

[b] Standard error of calibration ($\sqrt{\text{mean square error}}$) from the least squares regression of known forage quality values on NIRS values.

[c] Standard deviation of known forage quality values.

[d] 1 = 1st derivative; 2 = 2nd derivative treatment of NIR spectra.

[e] Wavelengths needed for best prediction equation.

[f] DM = dry matter.

[g] CP = crude protein.

[h] ADF = acid detergent fiber.

[i] NDF = neutral detergent fiber.

Table 6.—*Validation data for NIRS equations developed from information given in table 5*

Forage quality assay	N	r^{2a}	SEA[b]	Mean NIRS	Mean Chem	SD[c] NIRS	SD[c] Chem
Hay							
DM	29	0.89	0.74	94.6	94.6	2.1	2.2
CP	22	.98	.58	17.8	17.6	4.0	4.1
ADF	22	.89	1.32	36.5	36.9	3.7	4.0
ADF-CP	26	.82	.40	2.9	2.9	.9	.8
NDF	22	.91	2.59	52.1	52.3	7.3	8.4
Hay crop silage							
DM	25	.95	.61	93.7	93.6	2.2	2.5
CP	25	.90	1.08	17.5	17.0	3.2	3.4
ADF	25	.86	1.88	38.3	38.0	5.1	4.4
ADF-CP	27	.62	.82	3.5	3.6	1.3	.8
NDF	22	.77	2.83	49.5	50.4	5.0	5.9
Corn silage							
DM	25	.95	.50	94.4	94.3	2.1	2.2
CP	20	.72	.50	9.1	8.8	.9	.9
ADF	20	.95	1.43	26.4	26.5	6.0	5.5
NDF	20	.93	2.19	43.6	44.5	7.2	8.0
High moisture corn							
DM	31	.84	.84	92.9	93.0	1.9	2.1
CP	34	.82	.36	10.0	9.9	.7	.9
ADF	34	.98	.57	5.0	5.0	3.6	3.9
NDF	34	.97	1.23	11.9	12.0	6.6	7.0

[a] Squared simple correlation of NIRS analyzed values vs. known forage quality values from conventional laboratory assays.

[b] Standard error of analysis by NIRS.

[c] Standard deviation of laboratory forage quality values.

Table 7.—*Calibration statistics for 4 spectra files*

File No.[a]	CP (%) N	CP (%) Mean	CP (%) SEC	CP (%) R^2	ADF (%) N	ADF (%) Mean	ADF (%) SEC	ADF (%) R^2
1	95	14.6	0.49	0.98	95	49.0	1.12	0.94
2	110	17.1	.41	.99	110	37.5	1.20	.94
3	205	15.3	.44	.99	205	38.2	1.46	.91
4	317	15.4	.71	.98				

[a] 1 = spectra file from Merkle Laboratory, PA; 2 = spectra file from St. Paul, MN; 3 = spectra file of national hay equation from University Park, PA; 4 = files 1, 2 transferred to University Park and combined with spectra for 113 samples from the national hay file.

In situations where the master instrument cannot afford to support satellites with backup chemical analyses, spectra transfer is necessary. Spectra transfer is more difficult than equation transfer. However, spectra transfer provides the advantage of local equation update, provided the local NIRS laboratory has access to a well-controlled chemical analysis laboratory that matches the accuracy of the master instrument's laboratory.

Standardization of NIRS Analyses

The importance of selecting calibration samples that represent the population to be analyzed plus the importance of accurate chemical analyses in calibration have been discussed. Also important are instrumentation, computer selection, and computer software (see discussion of these topics in earlier chapters) because they affect the accuracy of the NIRS analyses. For a new technology to be successful in the field, it must be better than existing methods. NIRS is far superior to conventional chemical analyses in speed and accessibility to farmers, but its accuracy cannot be better than the chemical analyses that support it. Therefore, accurate, precise chemistry as well as proper calibration and instrumentation must be used to introduce this technology. In addition, the private sector is utilizing various instrumentation with various calibration methods; so NIRS approaches must be standardized, and the technology's accuracy documented.

Several Extension NIRS projects have the primary objective of transferring NIRS technology to private and public forage testing laboratories. Initial attempts suggest it will also be necessary to standardize chemical analyses to effectively transfer this technology.

The American Forage and Grassland Council (AFGC) and the National Hay Association (NHA) have formed an umbrella organization called the National Alfalfa Quality Testing Association to standardize tests for DM, CP, ADF, and EDDM (estimated digestible dry matter) on alfalfa hay nationwide. This organization will allow any laboratories testing alfalfa hay within the United States to voluntarily check their analyses, including both chemical and NIRS analyses. However, a standardization procedure is needed that will allow a broader range of forage crops to be included so that NIRS can be effectively transferred for public or private use. Extension programs could provide this standardization as part of their educational work.

Users of NIRS Information and Technology

NIRS technology has the potential of becoming the most important means for ruminant livestock producers to solve a long recognized problem: high feed costs. Within this problem, specific concerns that require a better solution than that offered by chemical analysis are (1) long lag times between forage sampling and receipt of forage test results on the farm, which render many conventional test results inapplicable to forages actually being fed when rations are balanced; (2) inability to test hays at the site of purchase; and (3) lack of sufficient forage tests on a given farm to allow nutritional improvements to be made for ruminant livestock.

Rapid onsite NIRS analyses via a mobile unit can facilitate improvement in forage quality and profitability for specific ruminant livestock producers. Adams (*1984*) states that in Pennsylvania, ruminant livestock managers cooperating in their NIRS project are improving their herd performance by quickly adjusting ration formulations according to frequent NIRS analyses. Wisconsin's NIRS mobile unit and soil and plant testing laboratory have increased their volume of samples analyzed since NIRS was introduced in lieu of chemical analysis (Rohweder *1984*).

Public Sector

The primary focus of the public sector's use of NIRS is education; however, many States choose to include a forage testing laboratory service for farmers in conjunction with educational programs. States that have forage testing laboratories and satellite or branch research stations are well organized to implement NIRS networks. Also, area Extension-research centers, area Extenstion offices, or counties with large numbers of ruminant livestock populations are potential users of NIRS instruments. However, many of these locations would require equation transfer.

In addition, nonprofit corporations, for example, State Dairy Herd Improvement Associations (DHIA), are good candidates for utilizing NIRS because they have central laboratories, a field staff for sample collection, and a centralized computer information system for rapid return of test results.

Private Sector

NIRS technology is being utilized in commercial forage testing laboratories and to a limited extent in feed or crop consultant organizations with mobile NIRS units. Its use by major feed manufacturers has been slower

than that in commercial forage testing laboratories, but growth is occurring in both. Equation calibration is being supplied by individual companies or by private instrument manufacturers. To date the private sector's use of instruments in a network has been limited. Commercial testing laboratories represent a wide range in volume of business, especially concerning the proportion of the total laboratory business that is derived from testing forage crops. NIRS has the potential to be used by the private sector for product quality control and education or trouble shooting. However, there is little cooperation between NIRS instrument manufacturers concerning equation calibration methods. This places more demand for educational support on Extension services to facilitate transfer of NIRS technology, especially in the laboratory standardization areas.

Education of Users

Success of educational programs about NIRS technology is crucial to its transfer to ultimate users. The State Universities in Wisconsin, Pennsylvania, and Minnesota have been asked to provide more information on how to interpret forage test results after initiation of NIRS analysis projects. Rohweder (*1984*) reported that use of the mobile NIRS unit at hay auctions throughout Wisconsin in 1983 increased hay prices by $20 per ton. In addition, hay price correlated best with relative feed value. Insufficient funding of Extension projects, the lack of laboratory standardization, and limited cooperation between the private and public sectors have delayed transfer of NIRS technology. Transfer of this technology has also been hampered by a lack of uniformity in chemical analyses.

NIRS users must know the limitations of calibration, instrumentation, computer selection, and software selection to accuracy of analysis. Inaccurate NIRS analysis can negate the contribution this technology can make to improving forage crop utilization.

NIRS has been effectively used to analyze farm grown forages in Pennsylvania, Florida, Vermont, Wisconsin, Oregon, Illinois, and Minnesota. Utah is also beginning an NIRS project on hay marketing. Minnesota utilizes NIRS to rapidly analyze alfalfa hay in windrows in its alfalfa grower program; this program would not be practical with long delays associated with conventional chemical analyses (Martin et al. *1984*). NIRS can be used in educational activities such as barn meetings, animal nutrition clinics, forage sampling clinics, and hay and silage shows. However, its most unique capability—that of providing nutritional information about forages before harvest or storage—has yet to be well used. If this capability is fully exploited, forage quality control will be greatly improved because NIRS analyses will identify those standing forage crops that can provide the quality desired for a specific livestock performance and will also suggest losses in forage quality arising from harvest and storage.

References

Adams, Richard S.
1984 Use of NIR, a rapid computerized forage testing method. *In* Proc. 9th Annual Minnesota Forage Day. p. 15-19. Minnesota Forage and Grassl. Coun., St. Paul.

Martin, N. P., D. A. Schriever, G. C. Marten, and F. R. Ehle.
1984 Use of near infrared reflectance spectroscopy (NIRS) to analyze forage quality of alfalfa grower samples. *In* Proc. 1984 Forage and Grassl. Conf, p. 223-227. American Forage and Grassl. Coun., Lexington, KY.

Rohweder, Dwayne A.
1984 Wisconsin's NIR and hay marketing projects—a progress report. *In* Proc. Wisconsin Forage Coun. 8th Forage Production and Use Symposium, p. 1-12. Wisconsin Forage Coun., Madison.

Templeton, W. E., Jr., J. S. Shenk, K. H. Norris, and others.
1983 Forage analysis with near-infrared reflectance spectroscopy: status and outline of national research project. *In* Proc. 14th Int. Grassl. Cong. Lexington, KY, p. 528-531. Westview Press, Boulder, CO.

NIRS Technology Transfer to Industry

J. S. Shenk

NIRS analysis for quality is only one component of an integrated system needed by the agricultural industry. The advantages of this type of analysis have been discussed in other parts of this handbook. The major point to be emphasized here is that it is computerized. Other important components of the complete system include forage and feed production, formulation, marketing, and farm management.

Feed quality analyses are important for quantifying plant-animal relationships. This quantification is the key to successful management and pricing. An example is the usefulness of NIRS analyses in hay marketing and ration formulation.

NIRS was proposed to help solve the analytical problem in hay marketing. A study was conducted in Pennsylvania to test the accuracy and feasibility of using NIRS in hay marketing (Shenk *1981*). Calculated errors of analysis for crude protein, ADF, NDF, and moisture were of the same relative size as the error of taking duplicate samples from a single bale of hay. This study also revealed that the price farmers paid for hay based on visual examination correlated poorly with quality as determined by analysis.

In a second study, NIRS equipment was installed in a mobile van and taken to the hay markets over a period of 6 weeks. This study demonstrated that hay could be analyzed onsite without disrupting the sale. The study is currently being expanded, and the results will be made available to farmers through Extension programs in Pennsylvania, Wisconsin, Minnesota, and elsewhere.

Not only do NIRS analyses need to be made at hay auctions, but they also need to be used in electronic marketing, which allows buyers and sellers to meet only by computer terminal. In this market, other information as to species, mold, and foreign material will need to be supplied to support the analyses.

NIRS analyses for formulation of rations are available at many places. The computers used for the NIRS analyses can also serve to balance the rations for dairy cows. These NIRS-computer systems, such as the ones developed by Pacific Scientific of Silver Springs, MD, are especially designed for the dairy industry. The software includes calibrations for the major forages and grains, library files for feed products not included in the calibration, business software for accounting purposes, client files for regular customers, ration formulation programs, and NIRS support programs.

Although the required software is sophisticated and complex, it is user friendly and flexible for the dairy consultant. Even though the NIRS analysis of the samples is very accurate and the computer cow model extremely comprehensive, the data provided by the dairy farmer in an interactive mode with the computer and dairy consultant are often very important aspects of the procedure.

Reference

Shenk, J .S.
1981 Using the near infrared analysis system to predict hay quality. Eleventh California Alfalfa Symposium.

Future Applications for NIRS

Feed Ingredients
F. E. Barton II

One future application of NIRS which is going to be quite important is the analysis of feed ingredients. Most of the work by the Network to date has concentrated on forages, especially in relation to plant breeding, forage management, and conserved forage evaluation studies. Not enough work has been done on total feeding regime. Forage is often part of a total diet, and the principal reason for analyzing its individual components is to ascertain their quality for ration balancing. The ingredients as well as the total ration must be analyzed for various compositional factors as well as quality factors. Another important aspect of component analysis is to provide the computer with information that will enable it to automatically select the appropriate ingredient equation to use for analyzing a mixed ration.

One caveat concerns using theoretical rather than practical information acquired by such a procedure. The analysis must reveal essential information about the feed ingredient and the ration. There are qualitative as well as quantitative data which should be considered. Many times a required result may be nominal rather than actual. A yes or no answer may yield all the required information. To know that an ingredient exceeds or fails to meet a threshold of acceptability is as important as a quantitative analysis. In essence, the instrument should respond as an analyst would when given specific data; that is, it should decide if the value determined is adequate.

Instruments that include expert systems such as this will require more sophisticated and flexible software as well as computers of enhanced capabilities. It is certainly feasible to expect more than one type of analytical instrument (that is, in addition to or instead of NIRS) to be tied to the computer information system. The future system for which we are producing a prototype technology will be of two types. The first is the large integrated system mentioned above. The second will be small, possibly portable, single use, dedicated instruments for a specific analysis or commodity. These two types of instruments will be able to handle the requirements of the scientist, producer, and regulatory groups for the analyses of feeds and feed components.

Mixed Feeds
S. M. Abrams

Application of NIRS for analysis of concentrate mixtures would permit feed mill operators to monitor the accuracy of feed formulations and allow livestock feeders to evaluate purchased concentrates and attain increased accuracy in diet formulation. Analysis of complete rations would indicate whether they are being prepared correctly.

Preliminary attempts to evaluate mixed feeds at University Park, PA, yielded poor results. It was concluded that this was due to the great heterogeneity of mixed feeds. Subsequently, an experiment was conducted to determine whether NIRS could measure the protein content of concentrates differing in the proportion of six ingredients: corn, oats, soybean meal, wheat bran, molasses, and dicalcium phosphate. The ingredients were ground in a Wiley mill (1-mm screen) and mixed in various ratios to provide 68 different concentrates mixtures (table 8) ranging from pure samples of each ingredient to mixtures of 5 ingredients. A calibration equation using linear multiple regression (the standard method) had a standard error that was twice as high as is usually observed in the analysis of protein in forages, and removal of mixtures containing particular ingredients that may have interfered with analysis did not result in any marked improvement (table 9). However, the application of nonlinear multiple regression equations dramatically improved the accuracy of calibration, reducing the standard error of calibration by approximately 50 percent.

These samples were then subjected to further grinding to ascertain whether particle size effects were increasing the variability of analytical values. This indeed proved to be the case (table 10). Use of the cyclone mill, which yields a finer and more uniform particle size, greatly reduced the standard error of calibration. There also appeared to be a slight advantage in using nonlinear regression in combination with the cyclone mill.

Table 8.—*Formulation of concentrate mixtures*

Number of mixtures	Number of ingredients per mixture	Ingredient ratios
6	1	100%
15	2	50%:50%
20	3	33%:33%:33%
15	4	25%:25%:25%:25%
6	5	20%:20%:20%:20%:20%
6	5	68%:8%:8%:8%:8%

Only tentative conclusions can be made on the basis of this artifically synthesized calibration set. However, it does appear that the use of finely ground material is essential when analyzing mixed feeds and that more complex regression techniques may further reduce the variation in analytical values obtained with NIRS.

Table 9.—*NIRS analysis of crude protein in mixed feeds utilizing linear and nonlinear multiple regression*

Calibration set	Linear SEC[a]	R^2	Non-linear SEC	R^2
All 68 mixtures	1.42	0.980	0.67	0.995
All mixtures except those containing dicalcium phosphate	1.48	.978	.75	.994
All mixtures except those containing molasses	1.41	.980	.77	.994
All mixtures except single ingredient mixtures	1.43	.976	.67	.995

[a] Standard error of calibration.

Table 10.—*Effect of particle size of mixed feeds on protein analysis by NIRS*

Mill (screen size)	Linear SEC[a]	R^2	Non-linear SEC	R^2
Wiley (1 mm)	1.42	0.980	0.67	0.995
Cyclone (1 mm)	.63	.996	.46	.998
Cyclone (.5 mm)	.60	.996	.49	.998

[a] Standard error of calibration.

High-Moisture Feedstuffs, Including Silage

T. H. Blosser

Situation

Feedstuffs containing appreciable amounts of water are an important source of nutrients for ruminants. Large tonnages of forage, preserved as silage, are stored in silos—towers, bunkers, trenches, stacks, and plastic bags. In many dairy areas of the United States, corn silage is of great importance; and in the more humid regions, hay crop nutrients can be stored with considerably less loss in the form of silage than as hay. In recent years, high moisture cereal grains have become popular feeds; and in some areas, byproducts of the brewing and distilling industries are fed in a wet form. Byproducts of the fruit and vegetable processing industries are frequently used as feeds by dairy and beef producers. Cull fruits and vegetables are also readily available as cattle feeds in certain sections of the United States.

Some of the same chemical characteristics that are important in determining digestibility and dry matter intake in dry feeds are also important in high moisture feeds. The amount and degradability or solubility of protein and amount and composition of the fiber components (hemicellullose, cellulose, lignin, and so forth) have an important effect on the amount and rate of digestion of both high moisture and dry feeds. When high moisture feeds are submitted to a laboratory for analysis, the most common procedure is to oven-dry the samples, grind them, and submit them to chemical analysis in the dried form. The same equations that are used for predicting digestibility and nutritive value of dry feeds are also used for high moisture feeds.

Analysis of silage to estimate its nutritive quality offers special problems (Van Soest *1982*). Often, an otherwise digestible product is associated with poor intake and animal efficiency. Poor quality silages are usually poorly consumed relative to hay of comparable digestibility. When fresh green forages are ensiled, they undergo extensive fermentative changes which can markedly affect their nutritive value. Soluble carbohydrates which are found in the growing plant are fermented to short chain acids—particularly lactic, acetic, and butyric. Protein and other nitrogenous components are degraded to smaller molecules during fermentation; and varying amounts of nitrogen-containing compounds of low molecular weight, such as ammonia, amides, and amines, are produced (McDonald *1981*). Buchanan-Smith (*1983*) found that amine-nitrogen levels were most closely and negatively associated with intake of alfalfa silage dry matter by lactating cows. The fermentative process generates considerable heat; and, if the

temperature of the ensiled mass is sufficiently high, reactions occur which render the protein relatively insoluble and resistant to enzymatic breakdown in the digestive tract (Thomas et al. *1982*). Langston et al. (*1958*) described good quality silages as containing large amounts of lactic acid and poor quality silages as being high in butyric acid.

Because of these characteristics, the conventional chemical analyses may present a misleading picture of the nutritive value of high moisture feedstuffs which have undergone fermentation. For example, because many of the fermentative end products are volatile, oven-drying gives an inaccurate picture of the true moisture content of these materials. The degree of error varies with the level and type of volatiles contained. Chemical analyses (that is, crude protein and acid and neutral detergent fiber), which are commonly used to evaluate dry forages, give virtually no information as to the type of fermentation which has occurred in a high moisture feedstuff.

Potential Role of NIRS in Analysis of High Moisture Feedstuffs

Waldo and Jorgensen (*1981*) stated that before NIRS can be used generally in forage and ruminant feed evaluation, it must be adapted to analyses of undried silages and high moisture grains. An immediate and obvious advantage of using NIRS directly on wet materials is to save time and, in the case of fermented materials like silages, increase accuracy.

Problems exist in analyzing samples, such as silages, which are high in moisture. One problem is the difficulty in preparing samples of sufficient homogeneity so that sampling is accurate. Dry samples are easily ground to a very fine state of subdivision through equipment such as a Wiley mill or a cyclone-type mill, but preparing a finely ground wet sample is more difficult. If these grinding problems can be solved effectively and inexpensively, the use of NIRS to analyze wet materials directly will increase substantially. Another problem with the direct NIRS analysis of wet materials is the presence in the NIR spectra of several strong absorption bands caused by water at 1480 and 1930 nm; and in peaks occurring above 2500 nm, a shoulder affects spectral values as low as 2300 nm. These absorption bands limit the usefulness of NIRS for high moisture feedstuffs in the spectral regions indicated. However, NIRS has been used successfully with a wide variety of high moisture materials in spite of this problem.

Abrams et al. (*1983*) produced 73 silages in small laboratory silos, analyzed the wet samples by conventional chemical methods, and also examined them in the wet state with NIRS. Correlations between chemical methods and NIRS predictions were greater than 0.90 for dry matter, nitrogen, and insoluble nitrogen but less than 0.50 for acetate, butyrate, and lactate.

In conclusion, near infrared reflectance spectroscopy gives promise of providing a technique for more rapidly and accurately determining certain chemical components in high moisture feedstuffs. However, additional research is needed to determine how to prepare wet samples for NIRS examination and to specifically identify those chemical compounds that are important to feedstuff quality and can be measured satisfactorily with NIRS.

References

Abrams, S. M., J. S. Shenk, and M. O. Westerhaus.
1983 Capability of near infrared reflectance spectroscopy (NIR) to determine silage quality. J. Dairy Sci. 66 (S.1):183.

Buchanan-Smith, J. G.
1983 Constituents in silage associated with feed intake reduction. J. Dairy Sci. 66 (S.1):184.

Langston, C. W., H. Irvin, C. H. Gordon, and others.
1958 Microbiology and chemistry of grass silage. U.S. Dep. Agri. Tech. Bull. 1187. 73 p.

McDonald, P.
1981 The biochemistry of silage. John Wiley and Sons Ltd., New York.

Van Soest, P. J.
1982 Nutritional ecology of the ruminant. 374 p. O and B Books, Inc., Corvallis, OR.

Waldo, D. R., and N. A. Jorgensen.
1981 Forages for high animal production: nutritional factors and effects of conservation. J. Dairy Sci. 64:1207-1229.

Thomas, J. W., Y. Yu, T. Middleton, and C. Stallings.
1982 Estimations of protein damage. *In* Protein requirements for cattle: symposium, MP-109, p. 81-98. Oklahoma State University.

Other Forage and Feed Nutrients

D. H. Clark

There are many additional types of feed nutrient analyses that might be made by NIRS, for example, analyses of fat- and water-soluble vitamins and fat composition of individual and complete feeds. NIRS analyses of fat and amino acid compositions of cereals have been successful; however, there are many other feedstuffs for which amino acid analysis should be investigated.

Macro minerals have been analyzed fairly accurately by NIRS. Analyses of trace minerals should be attempted even though they may be unsuccessful due to the low levels present. NIRS analyses for metabolizable or escape (bypass) proteins are additional areas that need research.

NIRS may be a rapid and reliable tool for determining the levels of constituents, such as alkaloids and other toxins, that can be harmful to livestock when they exceed specific levels in feedstuffs. Use of NIRS to detect additives (for example, ammonium sulfate and urea) could be important in feed marketing.

In fact, one of the most important potential uses of NIRS is to detect and identify a feed ingredient prior to, or concurrently with, making a quality measurement. This capability is desirable since mixed feeds with the same levels of protein, fiber, and other constituents may derive these constituents from different sources. For example, a 12 percent protein diet supplemented with soybean meal may be different from a 12 percent protein diet supplemented with cottonseed meal. In this example, the type of constituent in a feed may be as important as the quantity of a particular constituent present. Other possible uses of NIRS is to provide nominal data for monitoring the mixing of feed ingredients and maintenance of quality control. Inherent in quality control and identification is the need to determine moisture level on a continuing basis. Similar approaches have been applied to the flour-making process to control and improve process efficiency. These are but a few examples of potential new feedstuff analyses that should be attempted with NIRS.

NIRS Procedure In Technology Transfer

F. E. Barton II

For an NIRS method to be widely adopted, it must become an Official Method. Since agricultural products are bought, sold, and priced according to quality guidelines, those guidelines must be ascertainable by both parties to any transaction. The original Association of Official Agricultural Chemists, which was formed in Atlanta, GA, in 1884, has become the Association of Official Analytical Chemists (AOAC). The purpose of this century-old group is to adopt Official Methods for the analysis of agricultural products and chemicals. The Official Methods are used by regulatory analysts, research scientists, industrial scientists, and procurement specialists, who list AOAC methods in contract specifications. Currently, AOAC has certified three chemical methods for which the NIRS Network routinely develops analysis equations. They are crude protein (CP), acid detergent fiber (ADF), and dry matter (DM). The Network intends to apply for certification of an NIRS method based on the evaluation of CP, ADF, and DM.

AOAC adopts Official Methods on the basis of collaborative studies. The minimum collaborative study for a new method is generally considered to include 5 laboratories and 3 sample pairs to provide a total of 30 values for statistical treatment. The Network to date has completed two collaborative studies which qualify. The initial study involved 30 samples (5 of which were blind duplicates) and 6 laboratories, followed by an experiment with the same 6 laboratories to evaluate equation transfer with 60 samples. A third study is underway to further evaluate equation transfer at the 6 locations and to validate the transferred equations with 10 new samples. This experiment represents a third valid collaborative study.

In order to completely cover all aspects of the method, one more study should be completed. In this study, a set of equations developed by a single laboratory for a particular forage will be provided to all collaborators. A set of 60 samples will also be provided by this laboratory. The 60 samples will be scanned and then analyzed for CP, ADF, and DM by the transferred equations and by conventional chemistry in all cooperating laboratories. The chemical values will be used in addition to the 60 values used as the original calibration set to develop a new series of equations in all laboratories. An additional 20 samples will be provided by the original laboratory and used for a secondary validation of the original transferred equations as well as for validation of the new calibrations by all laboratories. These validation samples will also be analyzed by conventional chemistry to provide a further check (a second set of samples) for between-laboratory errors.

If these studies go as planned, we can expect the NIRS method to become an Interim Official First Action in 1985 or early 1986 and become an Official Method in 1987. Any subsequent analyses can append the Official Method without interim action. Appendment would require a collaborative study utilizing additional methods to be added to those that are part of the Official Method.

Laboratories in NIRS Technology Transfer

F. E. Barton II

It is not the intent or an objective of the NIRS Network to certify laboratories. That is the duty of the professional organizations and States which find it necessary to do so. The Network does want to establish an Official NIRS Method and will seek to do so through the AOAC. Basically, for any laboratory to become certified by some official sanctioning group, it must meet three criteria. First, its conventional chemical methods must be certified to meet certain standards, since they are the basis for the NIRS analysis. Second, the laboratory's instrument must be certified to meet certain specifications for signal-to-noise (S/N) ratio, wavelength accuracy and repeatability, and appropriate wavelength regions. The latter is very important since many instruments will not be scanning monochromators. As part of the second criterion, the instrument must accept appropriate software that has adequate programs for selection of spectra, evaluation of samples, and calibration if possible. The third criterion is the certification of calibration equations. These equations when applied to the spectral data yield the analytical results. The calibration may or may not be the same for any two instruments but should give the correct result for the same sample. When instruments have the capability to generate calibrations (usually the scanning monochromator or scanning filter systems), the equations will have to be validated by another instrument or with samples of known composition. When the instruments do not have calibration capabilities, external calibrations will be required that are certified by some agency or organization. Since some calibrations are commercially available and transferable to any instrument, they also must be certified and the limitations of the equations delineated for each analysis, feed ingredient, and range of values for which the equation will perform best.

Professional groups and State testing laboratories must then provide validation samples to certified laboratories on a regular basis. The procedure would require that a master instrument be set up to hold the master equation that serves as the final validation. While this is sound in theory, in practice there will never be a truly master instrument. Instruments break and need repair. At these occurrences, they change and their original equations may become invalid.

Spectra and equation transfer routines, along with transfer samples, allow the master laboratory with the master instrument to reestablish its master equation, either directly through transfer of its own original instrument's equations to its now changed instrument or through transfer from a second instrument.

These proposals are offered as a guide and are intended to prompt more questions from the readers and users of this manual and the NIRS technique that they are intended to answer. NIRS analysis and chemometric methods are new concepts of analysis that promise speed and economy previously unknown; but they require refinement to be truly usable tools for agriculture. It is hoped that the reader will be able to apply the NIRS technique with a clearer understanding of the procedure. All the NIRS concepts mentioned here are discussed in detail earlier in the handbook, and it is essential to understand the concepts before applying them to a complete analysis protocol.

Appendixes

1. History of the National NIRS Forage Research Project Network

G. C. Marten and W. C. Templeton, Jr.

Developing a fast, accurate, and precise laboratory procedure to estimate forage quality has been a major research goal for many years. The potential of using NIRS to estimate forage quality was first suggested in April 1975 at a workshop, cosponsored by USDA-ARS and the American Forage and Grassland Council, in Beltsville, MD.

Cooperative Agreement Between USDA-ARS and The Pennsylvania State University, 1975, and Progress to 1977

R. F. Barnes, at the U.S. Regional Pasture Research Laboratory (USRPRL) in University Park, PA, was instrumental in initiating the cooperative research that provided the first evidence that NIRS could measure forage quality. This research was stimulated by the earlier successful use of NIRS for evaluating the quality of grain and oilseed crops at the Instrumentation Research Laboratory (K. H. Norris, Chief) of the Beltsville Agricultural Research Center. USDA-ARS established a Specific Cooperative Agreement with The Pennsylvania State University (PSU) in June 1975 to investigate in depth the use of NIRS for evaluating the quality of forages and other feedstuffs. The original project leaders were R. F. Barnes (USDA-ARS), J. S. Shenk (PSU), and B. R. Baumgardt (PSU).

A scanning monochromator NIRS instrument was designed and assembled for the project by Lemont Scientific, Inc., State College, PA. Encouraging progress was made in rapidly evaluating the feeding quality of forages. When W. C. Templeton, Jr., became director of the USRPRL in 1977, he reviewed the status of NIRS research in Pennsylvania and other locations. He called a meeting in October 1977 on "Statistical Considerations in Forage Evaluation by Infrared Reflectance Spectroscopy." Another meeting was held in December 1977, involving five chemists and a statistician from the Eastern Regional Research Center, Philadelphia, a plant breeder from Beltsville, and four scientists from University Park. Discussions resulted in a consensus that the potential for using NIRS in forage quality and animal nutrition research, ration formulation, hay grading, hay marketing, and Extension programs warranted a major research thrust.

National NIRS Forage Research Project Proposal and Formation, 1978 to 1980

In 1978, Templeton and Shenk drafted a research proposal and budget for a cooperative NIRS project to evaluate forage crop quality. To be established were a central laboratory at University Park, PA, and satellite laboratories at strategic locations in the Nation. This proposal was sent to USDA-ARS National Program Staff scientist Barnes, who then called meetings to establish an information system for NIRS, criteria for selecting cooperating laboratories, coordination of the research proposed, and objectives for the National NIRS Forage Research Project Network. Primary participants in this total effort were Barnes, F. E. Barton II, D. Burdick, L. L. Jansen, G. C. Marten, W. H. Martinez, Norris, J. B. Powell, H. Puterbaugh, T. Schatzki, Shenk, and Templeton.

In February 1979, a formal request was submitted to USDA-ARS for complete or partial funding of NIRS equipment to be used at Athens, GA; Beltsville, MD; St. Paul, MN; El Reno, OK; University Park, PA; and Logan, UT. These locations had been selected to be part of the National NIRS Forage Research Project Network. The proposal was approved for funding in April 1979. Templeton was named as the first project coordinator for the Network.

The original proposal for the National NIRS Forage Research Project Network included the following five objectives:

1. To develop and test computer programs which provide continuing advances in data processing and mathematical treatment of infrared data to maximize prediction accuracy;

2. To further define and measure plant, environmental, and other factors contributing to variation in infrared prediction of chemical composition and animal response;

3. To relate chemical and physical properties of forages to infrared reflectance spectral properties;

4. To test the usefulness of infrared reflectance in forage-breeding, forage-management, and animal-utilization research programs; and

5. To produce, analyze, assemble, and maintain selected forage samples in a reference library for use in infrared instrument calibration and other forage evaluation studies.

Listed according to location, the original primary research personnel who contributed to these objectives were:

Location	*Personnel*
Athens, GA	F. E. Barton II W. R. Windham
Beltsville, MD	J. H. Elgin, Jr. K. H. Norris
St. Paul, MN	J. L. Halgerson G. C. Marten
El Reno, OK	S. W. Coleman F. P. Horn
University Park, PA	M. R. Hoover J. S. Shenk W. C. Templeton, Jr. M. O. Westerhaus
Logan, UT	M. J. Anderson

After discussion of instrumentation needs, with primary inputs from Norris and Shenk, USDA-ARS circulated to prospective suppliers in June 1979 a request for proposals to supply six computerized, high-precision, near infrared reflectance spectrophotometers. The proposal of Neotec Corporation, Silver Spring, MD, was eventually accepted, and delivery of equipment was initiated in November 1979 and completed in January 1980.

In 1980 activity was intense, even though instrument delivery, familiarization, and some hardware problems allowed less than a full year's research time for some locations. A set of 30 forage samples was used to compare chemical and spectral assays within and among the 6 locations. Both cool-season and warm-season species were included. Each laboratory conducted standard laboratory analyses of the forages. In addition, each collected reflectance data for each of the samples on 3 consecutive days.

Average spectral curves from the six instruments were noticeably different in reflectance level and in amplitude of absorption bands, but a simple normalization of the data resulted in curves essentially alike except in the water regions. Differences in the normalized curves were caused primarily by differences in sample moisture at the six locations.

The errors associated with the NIRS analyses compared favorably with those obtained by routine chemical and biological procedures. Moreover, between-laboratory variations were no greater for NIRS than for the standard procedures.

Near the end of the year it was clear that a group meeting to discuss progress, operating problems, and future plans was essential. The first annual NIRS forage research network workshop was scheduled at the Southwestern Livestock and Forage Research Station, El Reno, OK, in February 1981.

Highlights of National NIRS Forage Research Project Network Reports of Progress and Revised Plans of Work, 1981 to 1984

First Workshop

In February 1981, the NIRS Network held its first annual workshop at El Reno. Eighteen scientists participated. This and the next two workshops were organized by local hosts and W. C. Templeton, Jr., who presided also. Besides local reports of research progress from the six Network locations, the workshops featured discussions whose topics included statistical results of NIRS and chemical analyses of forage samples distributed for collaborative research; NIRS terminology and interpretation of data output; experiences with software, including the need for new programs; calibration procedures; hardware problems, including need for rotating sample holders, scanning indicator lamps, line conditioners, and hardware troubleshooting; new NIRS applications to forage quality assessment; monochromator maintenance; and interest in NIRS technology for forage-quality analysis at several research stations not included in the Network.

Among successful or partially successful forage-quality measurements that could be made by the scanning monochromator instruments were those for small grain crops, corn stover, and cool-season perennial grasses in Minnesota; alfalfa in Utah; and tropical grasses in Georgia. Minnesota reported also that the same grinding procedure should be used for calibration samples as for analysis samples to ensure best results. Grinding forage samples through a 1-mm-mesh screen in a cyclone mill gave the best precision.

The collaborative study, in which 30 forage samples were subsampled for analysis in each Network laboratory, revealed that among-laboratory variations for the chemical analyses exceeded those for the NIRS analyses. Also, this study revealed that the instruments

(all built by the same manufacturer following the same specifications) differed in that the spectrum was vertically offset at different locations. The need for standardization and improved accuracy of chemical procedures was revealed to be the biggest problem in obtaining excellent NIRS calibration equations.

Summer 1981

Special reports of progress by late summer 1981 revealed several new accomplishments and problems.

Accomplishments. Forages that were newly and successfully analyzed for forage quality measurement by NIRS at the Network locations included birdsfoot trefoil-grass-weed mixtures (Minnesota); Old World bluestem and alfalfa (Oklahoma); crested wheatgrass (Utah); alfalfa and tall fescue (Maryland); and alfalfa and 60 farmer-hay samples (Pennsylvania). The Maryland laboratory developed (1) the derivative-ratio data treatment for NIRS spectra to minimize effects of sample temperature, particle size, and interfering absorbers and (2) the curve-fitting data treatment for NIRS spectra to provide a more robust calibration when an adequate number of samples is available.

The Pennsylvania laboratory (1) developed a complete software package for operating the Neotec 6100/PDP 1103 (1134) spectrocomputer system; (2) developed a mobile NIRS van which was driven over 34,000 miles to test utility of the equipment for assessing hay quality at markets and quality of forages used in dairy cattle rations; and (3) found evidence for the importance of including a range of calibration samples so that they truly represent samples for which quality values are to be analyzed by NIRS.

Problems. NIRS operating or maintenance problems reported by Network laboratories included (1) the need for lamp and encoder replacement and the occurrence of extensive computer problems by the end of the first year (Minnesota); (2) considerable hardware problems in startup and early operation (Oklahoma); (3) considerable difficulty with the computer, monochromator, and other hardware, including the need to repair the floppy-disk drive and to replace the lamp before the end of the first year (Utah). Georgia, Maryland, and Pennsylvania reported no significant problems during the first year of instrument operation.

Second Workshop

The second NIRS Network workshop was held at Athens, GA, in December 1981. Among the 19 participants were several researchers who had not been originally included as contributors to the objectives (D. Himmelsbach and J. Robertson of Athens, GA; R. Meyer of El Reno, OK; and C. Liebman of Ona, FL).

Minnesota reported continued success in NIRS measurement of forage quality in corn stover, cool-season grasses, and legume-grass mixtures, but they reported also the need to update previous analysis equations to best accommodate new samples grown in succeeding years. The NIRS analysis of reed canarygrass for neutral detergent fiber (NDF), acid detergent fiber (ADF), crude protein (CP), and in vitro dry matter digestibility (IVDMD) was routinely used to successfully screen thousands of genotypes in a plant breeding (heritability) project.

Oklahoma and Georgia both reported the merits of a semiautomated Fibertec system for quickly and precisely assaying NDF and ADF in calibration samples of grasses. Also, Georgia had begun to use the Karl Fischer method of measuring moisture in previously dried forage samples in order to develop NIRS calibration equations that could eventually be used to correct laboratory samples for varying moisture without redrying. Additional Georgia progress included use of high performance liquid chromatography for hemicellulose analysis; identification of botanical tissues in forages that had been extracted by fiber-assay procedures; use of ^{13}C nuclear magnetic resonance (NMR) for structure elucidation and quantitative measurements as a complement to NIRS; and calibration of the NIRS instrument to analyze CP and CP digestion from samples of the diet, abomasum, and feces.

Shenk (Pennsylvania) presented a tentative outline of procedures to be used for routine calibrations in NIRS analyses. It included selection of 70 samples from broad populations of forage samples, assessment of three math treatments, validation by use of 30 percent of the sample population (randomly selected), selection of the appropriate number of wavelengths, chemical reanalysis of "starred" samples (those that exceeded specificed multiples of the *H* statistic) in the computer printout, and periodic updating of equations with samples that expand the range of calibration.

Utah reported continued success in using NIRS for quality analyses of alfalfa hay and grasses in a breeding project. Norris (Maryland) reported that in preliminary tests six Network instruments *differed* in noise level, absolute reflectance, and sensitivity.

In addition to the above local reports of progress, the workshop featured discussions, which pointed out the need for rotating sample holders on a center pivot to

reduce error; the introduction of new comprehensive software with changes to eliminate noise output for bad scans, to remove restraints during incorporation of set functions, to add division terms that cancel particle size, temperature, and interfering absorptions, to improve prediction by 20 percent, and to measure noise difference between samples; the need for universal calibration equations for measuring crude protein and other forage-quality-constituents in all species; and the need to revise the original NIRS research objectives.

The discussion on the last need led to the following revised plan for the NIRS Forage Research Project Network:

Goal, Objectives, and Approaches of the National NIRS Forage Research Project Network (As Amended, 1982)

Goal:
Establish an improved near infrared reflectance analytical system for forage and feedstuff quality evaluation.

Objective 1:
Relate chemical and physical properties of forages to their infrared reflectance spectral properties and to their utilization by ruminants.

Approaches
a. Relate nuclear magnetic resonance and other analytical techniques to interpretation of NIRS results and chemical data,
b. Investigate sequential changes in NIRS spectra of feeds which occur in the process of digestion by ruminants, and
c. Determine effects of anatomical and other physical properties of forages on NIRS spectra.

Objective 2:
Test and validate NIRS for determination of forage quality.

Approaches
a. Define and measure plant, environmental, and other factors contributing to variations in infrared predictions of chemical composition and animal responses and
b. Develop and test computer programs which provide continuing advances in data processing and mathematical treatment of infrared data to maximize prediction accuracy.

Objective 3:
Establish standards for the conduct of NIRS analyses.

Approaches
a. Develop procedures for instrument standardization,
b. Develop protocol for calibration and verification of prediction equations, and
c. Develop procedures for sample preparation and presentation of samples to NIRS instruments.

Objective 4:
Establish and maintain a library of reference forage samples for use in NIRS instrument calibration.

Approaches
a. Produce/obtain samples with known chemical and nutritional properties,
b. Establish procedures and a facility for maintaining sample integrity, and
c. Catalog, provide, and distribute samples for use in forage-quality assessment.

Objective 5:
Facilitate transfer of NIRS technology.

Approaches
a. Establish levels of accuracy and precision of parameters required for use of NIRS in forage breeding and management,
b. Establish levels of accuracy and precision of parameters required for use of NIRS in animal-utilization research, and
c. Formulate guidelines for use of NIRS to assess the quality of forage and feed products in livestock feeding and the marketing system.

Third Workshop

The third NIRS Network workshop was held at University Park, PA, in February 1983. Forty seven participants attended, including a new research contributor (S. Abrams of USDA-ARS in Pennsylvania) and several Extension workers who planned to start technology transfer projects.

Georgia reported further progress in using NMR to help explain NIRS calibrations for quality constituents in forage and grain samples. They developed successful NIRS calibration equations for CP, NDF, and IVDMD. Oklahoma developed satisfactory NIRS equations for analyzing CP, NDF, and IVDMD in plant breeders' samples of eastern gamagrass, weeping lovegrass, and Old World bluestem. The new software improved calibrations. Utah reported an overheating problem in their

monochromator, which was eventually traced to a malfunctioning fan.

Minnesota reported distinct improvements in the accuracy of NIRS analyses for CP, NDF, and ADF by including year effects and harvest-within-year effects on reed canarygrass genotypes during calibration. However, corn stover equations remained robust without year effect additions. For quality analyses of birdsfoot trefoil-grass mixtures by NIRS, equations based on samples from 2 combined years were better than separate equations for each of the 2 years. Minnesota found that, compared to use of NDF concentration to predict the amount of legume, NIRS provided a superior prediction of percentage legume in birdsfoot trefoil-grass mixtures. Also, NIRS provided excellent analyses of CP, ADF, NDF, and IVDMD for alfalfa grown in a cutting-management experiment, and harvests by both sickle-bar and flail-chopper machines could be accommodated in the same equations. Minnesota further reported that generalized equations could be developed to analyze CP, ADF, NDF, IVDMD, and acid detergent lignin (ADL) in a wide range of legume samples (which included plant parts of four species) just as well as could separate equations for each species.

Pennsylvania reported substantial progress in four areas of NIRS research, including evaluation of current instruments on the market, work on new software programs, computer evaluation and instrument update, and calibration transfer among instruments. Also, improved corn-silage moisture analyses resulted when Karl Fischer rather than oven-dry values were used in the NIRS calibration. The Network scientists agreed to adopt the Karl Fischer procedure as standard for moisture assays of forages.

Norris distributed results of a comparison of the six NIRS Network instruments for measurements on polyethylene beads and on a ceramic standard. Several of the instruments differed considerably in the log (1/R) values, but those in Utah, Minnesota, and Pennsylvania were similar.

Shenk led a discussion of ways in which NIRS scanning instruments from different manufacturers operate. Westerhaus and Shenk compared old and new software programs and presented a comparative evaluation of computers. Among other extensive discussions were those concerning research-Extension cooperative relations, program development, and coordination for NIRS evaluation of forages; hay-quality evaluation and marketing; and silage-quality evaluation.

Fourth Workshop

The fourth NIRS Network workshop was held at Logan, UT, in November 1983 simultaneously with a meeting of the WRCC-48 Committee. Twenty-eight of about 50 persons attending joint sessions of the two groups were actively involved in most or all of the NIRS workshop (including new research contributors T. Blosser of Beltsville, MD; D. Clark of Logan, UT; and E. Redshaw of Edmonton, Alberta). Because of W. C. Templeton's retirement, this workshop was organized by P. V. Fonnesbeck (Utah host) and the new project coordinator, G. C. Marten, who presided.

Georgia reported comparisons of three software programs used to evaluate six forage-quality components of eastern gamagrass by NIRS. They obtained excellent calibrations for quality components of alfalfa in a breeding program. Substantial year and harvest effects for NIRS assessments of IVDMD and CP in bermudagrass were accounted for by combining files. Three forage-quality components in sericea lespedeza cultivars were satisfactorily measured by NIRS. Georgia also reported progress in using solid-state ^{13}C NMR to explain NIRS variables, after having solved considerable NMR equipment problems.

Maryland had engaged new personnel to initiate NIRS research with emphasis on forage sample drying and grinding methods and on high-moisture feedstuffs, including silage. Minnesota reported continued excellent NIRS measurement of forage-quality traits in reed canarygrass, four legume species, five grass species treated with mefluidide, alfalfa managed with diverse cutting treatments, plant parts of four grass species, mixtures of alfalfa and companion-crop small grain species, and mixed hays gathered by Extension workers throughout the State. Minnesota also reported excellent NIRS measurements of total nonstructural carbohydrates in alfalfa roots.

Major emphasis in Oklahoma was placed on NIRS calibrations to assess five quality traits in plant breeders' samples of warm-season perennial grasses. Also, Oklahoma reported progress in analyzing the nutrient compositions of feed and feed constituents. Because of the robustness of the data set, the heterogeneity of the proteins in the feedstuffs (feathers, blood meal, alfalfa, mixed commerical feeds, dog and cat foods, and pure lysine), and the range of protein concentrations (from 13 to 93 percent), the NIRS calibration for the nutrient composition equations had high R^2 values accompanied by high standard errors.

Utah successfully used NIRS to analyze soluble nitrogen and fibrous components in crested wheatgrass, carotene in alfalfa hay, and the chemical compositions of vegetables. However, Utah's anticipated progress was delayed by serious instrument problems for nearly a year (including a defective power supply board and excessive noise).

Pennsylvania submitted a report of substantial research progress in numerous areas, including NIRS equation transfer between instruments; spectra transfer between instruments; an instrument comparison and data processing study of 3 NIRS instruments; successful development of universal NIRS equations for analyzing CP, IVDMD, and DM from a population of 906 hay samples gathered from 31 States (joint effort by members of the Network and cooperating Extension specialists); an investigation of errors in NIRS analyses; analyses of concentrate mixtures, particle size effects, and data handling by nonlinear math; a new proposed NIRS software release which combined the best features of the CAL and BEST programs; and proposed hardware updates, including a floating point processor to speed up programs and a graphics terminal and plotter for each Network system.

Redshaw (Edmonton, Alberta) presented a detailed report of NIRS progress in forage-quality evaluation by the Agricultural Soil and Feed Testing Laboratory of the Alberta Department of Agriculture. Because their NIRS hardware and software were identical to those used by the members of the National NIRS Forage Project Network, their results could be directly compared to those of the Network. They reported development of NIRS calibration equations for selected assays, including moisture, CP, ADF, NDF, ADL, calcium, and phosphorus for forage grasses, forage legumes, cereal silages, forage grass and legume silages, grains, and barley straw.

Norris led a discussion of new developments for NIRS, including a proposal that a Halon-carbon mixture be used as a laboratory standard for NIRS; evidence that three NIRS instruments were very stable in wavelength measurement and within 1 nm in picking the wavelength for polystyrene; methods for and results of a noise test on different instruments; and bias error as a result of temperature changes in rooms housing instruments.

Martinez (USDA-ARS National Program Staff) led a discussion of new research opportunites and thrusts in NIRS as they relate to implementation of the ARS Strategic Plan. At least six subproblems of the Strategic Plan applied to the Network's research. A listing of commitments for future research within the Network revealed that at least 25 person-years would be applied to NIRS research on forages and other feedstuffs in 1984.

Minnesota and Pennsylvania Extension specialists reported progress in transferring NIRS technology for use by industry, farm producers, hay marketers, and others interested in assessing forage quality and ration formulation. Several Network members indicated their intent of cooperating with the Agricultural Extension Service to aid this technology transfer.

The need for an NIRS certifying laboratory was discussed in detail. The first step is to have an NIRS method certified as an Official Method for forage-quality analysis. The Network decided to approach the American Association of Analytical Chemists to accomplish this. Barton (Georgia) was designated as an "associate referee" to lead the certification effort. Martinez revealed that an organization such as The Smalley Committee of the American Oil Chemists' Society could be approached eventually to establish procedures for and the conduct of certifying individual laboratories.

After considering progress and completed objectives since the second workshop (last modification of objectives), the Network delegates agreed to a revised long-range plan for the National NIRS Research Project:

Long Range Plan
National NIRS Forage Research Project
(Revised November, 1983)

Goal:
Establish an improved analytical system for forage and feedstuff quality evaluation.

Objective/Strategy 1:
Test and validate NIRS for determination of forage quality.

Approaches
a. Define and measure plant, environmental, and other factors that contribute to variation in infrared analysis of chemical composition;
b. Define and measure plant, environmental, and other factors that contribute to variation in infrared prediction of animal response; and
c. Develop and test computer programs which provide continuing advances in data processing and mathematical treatment of infrared data to maximize analytical accuracy.

Objective/Strategy 2:
Test and validate NIRS for determination of quality of other feedstuffs, grain, concentrates, and mixed diets.

Approaches
a. Same as approach a of objective 1,
b. Same as approach b of objective 1, and
c. Same as approach c of objective 1.

Objective/Strategy 3:
Define the fundamental basis for the infrared reflectance spectral properties of feedstuffs as related to their chemical and physical properties and their utilization by ruminants.

Approaches
a. Investigate NMR and other analytical techniques to complement interpretation of NIRS results,
b. Investigate changes in NIR spectra of feedstuffs which occur in the process of digestion by ruminants, and
c. Determine effects of anatomical and other physical properties of forages on NIR spectra.

Objective/Strategy 4:
Facilitate transfer of NIRS technology.

Approaches
a. Establish levels of accuracy and precision of parameters required for use of NIRS in forage breeding and management,
b. Establish levels of accuracy and precision of parameters required for use of NIRS in animal utilization research,
c. Formulate guidelines for use of NIRS in quality assessment of forage,
d. Formulate guidelines for use of NIRS in quality assessment of feed products in livestock feeding and the marketing system, and
e. Facilitate transfer of NIRS technology via cooperation with Extension personnel.

Objective/Strategy 5:
Establish and maintain a library of reference feedstuff samples and spectra for use in NIRS instrument calibration and other forage-evaluation studies.

Approaches
a. Produce or otherwise obtain samples with known chemical and nutritional properties,
b. Establish procedures and a facility for maintaining integrity of stored samples,
c. Catalog and distribute samples as required, and
d. Provide for necessary laboratory analyses.

Objective/Strategy 6:
Establish standards for the conduct of NIRS analyses.

Approaches
a. Develop procedures for instrument standardization;
b. Develop protocols for calibration and verification of analytical equations, and
c. Develop procedures for sample preparation, including drying and presentation of samples to NIRS instruments.

Conference Calls and Troubleshooting

In addition to the annual workshops, communication among members of the Network, their advisors, and cooperators was greatly aided by frequent national conference calls sponsored by the USDA-ARS National Program Staff. Instrument and software troubleshooting was also accomplished by calls and laboratory visits among individuals in the Network; scientists at University Park, Pennsylvania (USRPRL and PSU), often served in the lead role.

The research required for preparing this handbook was conducted by genuine collaborative effort among numerous individuals of varying expertise both within and outside the Network over a period exceeding 5 years.

2. Considerations of Chemical Analyses

F. E. Barton II

Chemical Analyses

By definition a good analytical method is accurate, precise, rapid, and low cost; gives easily interpretable results; requires only simple sample preparation; and examines the structure in situ with no chance for sample preparation artifacts to alter the results. Chemometrics as a discipline within chemistry can be defined as the development and application of mathematical and statistical methods to extract useful chemical information from chemical measurements (Kowalski *1977*). For our purposes we include the extraction of compositional information from spectral curves in a chemometric method as described by Norris (*1983*a and *1983*b). The basic impetus for the development of predictive analyses is the increasing cost of performing laboratory analyses and the time required to obtain the results. We should first realize that we have been using chemometric methods for decades. Whenever we construct a standard curve from a series of standard solutions, assuming linearity from Beer's Law, and read the concentrations of unknowns directly from the scale, knowing only their absorbances, we have used a chemometric method. Using standard curves in analyses accomplished with ion-selective electrodes is another example. The great decrease in the cost of computers has made possible the technology to which this handbook is directed, NIRS. Because of the statistical routines, we can create the calibration equations (we really create standard curves in the computer) and then, simply from taking the spectrum, present an analytical result without having performed the analysis in question. NIRS is, therefore, a chemometric method. In fact, without the larger, current generation of microcomputers, many NIRS attempts would not be possible.

The advantages of the NIRS method for managing laboratory data can best be given by examples (Barton *1984*). In table 1 are the results of applying NIRS to a series of samples, five of which were blind duplicates. Data in columns labeled "SDP" and "SDA" show that for all forage constituents except crude protein, the method was more precise than the laboratory method in duplicate analyses. The numerical difference between SDP and SDA for protein probably is very close to instrumental error in the method. Overall, it would seem that precision as well as speed is improved when NIRS is used. It requires 5 weeks to run these analyses in the laboratory by gravimetric methods, but only 1 hour by NIRS.

The data in table 2 result from in vitro digestibility determinations on a set of Old World bluestem grasses. Samples 27 to 40 (with large negative biases, which indicate that the predicted values were larger than the measured values) had been chemically analyzed in triplicate, side by side, with quite acceptable standard deviations. When these samples were reanalyzed chemically, the new values agreed with the NIRS predicted values. This error could have been caused by a restriction in the automatic pipet or in the straining of the rumen inoculum, which lowered the number of rumen micro-organisms or lowered the activity of the inoculum due to residual detergent in the three racks of tubes. These samples represented killing the bacteria and retarding digestion. There was no way to check the standard laboratory results for an indeterminant error such as this. The results routinely would have been accepted, and an error of 10–30 percent would never have been detected. It is apparent in this example that ac-

Table 1.—*Calibration of NIRS instrument for a broad spectrum of samples*[a,b]

Analysis	N[c]	SEC[d]	R^{2e}	Repeatability[f]	SDP[g]	SDA[h]
IVDMD	3	1.74	0.92	0.16	0.37	1.72
CP	2	.72	.98	.02	.22	.09
ADF	3	1.6	.88	.08	.24	.28
ADL	3	.78	.87	.06	.12	.29

[a] Set of 30 samples included 13 bermudagrass, 5 alfalfa, 4 orchardgrass, and 1 timothy. Twenty were hays and 10 were fresh frozen and freeze dried.
[b] From Barton and Coleman (unpublished report).
[c] Number of wavelengths.
[d] SEC = standard error of calibration.
[e] R^2 = multiple correlation coefficient.
[f] Repeatibility = the measure of added signal to noise ratio for an additional term in the equation. If it increases with a new term, more noise than signal is being added and overfit is more of a problem.
[g] SDP = average standard deviation of prediction for 5 El Reno samples analyzed in duplicate by NIRS.
[h] SDA = average standard deviation of actual laboratory values for the 5 El Reno samples analyzed in duplicate.

Table 2.—*Actual in vitro dry matter digestibility (IVDMD) vs. predicted IVDMD for 60 forage samples*

Sample No.	Ave. res. size[a]	Bias
1-26	2.5	Pos. and neg.
27-36	10.0	All negative.
37-40	4.4	All negative.
41-60	1.7	Pos. and neg.

[a] Average residual size.

curacy can be improved through the use of chemometric procedures. Table 3, which concerns determinations of dry matter (DM), contains data showing both positive and negative biases. In the laboratory, trays were used to carry 20 dry crucibles at a time from the oven to the desiccator to the balance. The data reflect differential moisture absorption among the glass crucibles while they were stored on the trays in the desiccator. The chemometric procedure is very sensitive to laboratory practices which involve consistent errors and makes it quite easy to identify the sources of these otherwise difficult to detect errors.

Any discussion of chemical analyses and their relationship to NIRS analyses should begin with a list of premises for the analyst to consider:

1. The importance of good laboratory data cannot be overemphasized.

2. The NIRS results can be no better that the data used for calibration.

3. The analyst must thoroughly understand the chemical method to include the sources of error and what the method measures.

4. Often the analyst does not, in general, understand the chemical method and the errors associated with it.

5. Even if the analyst does know all the errors, they would likely prove to be larger than expected.

6. Precision is correctly defined as the closeness with which a result can be repeated (for example, deviation between spectral analyses).

7. Accuracy is correctly defined as the closeness with which the assay compares with the true answer.

8. For an empirical analytical procedure there is no way of determining the true answer.

Table 3.—*Actual dry matter (DM) vs. predicted DM for 60 forage samples*

Sample No.	Ave. res. size [a]	Bias
1-20	1.74	All negative.
21-40	0.42	0.0
41-60	1.78	All positive.

[a]Average residual size.

These premises are given as a basis for discussing empirical analytical procedures and what the data we generate mean and how we use them to interpret the NIRS results we will get from the instrumental procedure. An empirical method is defined by the conditions of the analysis rather than by the entity to be determined.

Fibrous materials traditionally have been analyzed by the Weende proximate analysis procedures as a means of estimating total digestible nutrients. In the proximate analysis procedures, percentage crude protein is expressed as 6.25 × percentage nitrogen from the Kjeldahl analysis, percentage fat by ether extraction, percentage crude fiber by alternate base (1.25 percent NaOH) and acid (1.25 percent HCl) treatments, and percentage ash by incineration (AOAC *1980*). These procedures continue to be the standard methods used by many State testing laboratories. These analyses are empirical. As such, there is the tacit assumption that the reagents or experimental conditions affect each sample in an identical manner. They are all gravimetric procedures, and the results are given in relative percentages. Two excellent reviews have been published which detail the status of gravimetric forage analyses (Moore and Mott *1973*, Martens and Russwurm *1983*). Since the molecular weight of a forage sample or any constituent therein cannot be determined, these percentages are the only way to quantitatively express compositional properties. The analyses are very dependent on sampling techniques, technician experience, and the environment in which the sample is analyzed. Finally, all of the procedures are destructive.

Van Soest published several procedures on detergent fiber analyses (Van Soest *1963*a, *1963*b; Van Soest and Wine *1967, 1968*) and a means of classifying forage fractions (Van Soest *1967, 1973*). The detergent analysis system defines the sample by separating it into two fibrous fractions—a neutral detergent fiber fraction (NDF), the residue obtained by extraction of the sample with buffered 2 percent sodium lauryl sulfate solution, and an acid detergent fiber fraction (ADF), the residue obtained by extraction of the sample with a boiling 2 percent solution of hexadecyltrimethylammonium bromide in 1.0 N sulfuric acid. NDF (Van Soest and Wine *1967*) is considered to be total fiber or cell walls, while ADF (Van Soest and Wine *1063*b) is the lignified portion of the plant cell walls. The ADF procedure primarily is used as a preparatory step for lignin determinations, either with 72 percent sulfuric acid (Van Soest *1963*b) or with permanganate (Van Soest and Wine *1968*). The detergent analysis procedures, ADF and NDF in particular, have been used with varying degrees of success to predict extent of digestion and

relative intake, respectively (Rohweder et al. *1978*). In recent years, these procedures have begun to replace the crude fiber procedure as the standard analyses for fiber in forages. The ADF procedure is now an AOAC approved analysis. The literature values for many forages in the above references are comparable to those in table 4, which were obtained by the detergent analyses procedures. However, the tacit assumption is that the reagents provide the same treatment in all samples regardless of species, environment of growth, and agricultural management practices.

Almost all State testing laboratories, commercial analytical laboratories, and research laboratories will at some time determine mineral content of a forage or feed material. The data obtained are used in ration formulation and balancing calculations. The particular analyses can be done by wet chemistry, atomic absorption spectroscopy, or other instrumental methods. It would be impossible to describe all the methods in this handbook, but the 13th edition of the manual "Methods of Analysis" (AOAC *1980*) should be the prime reference source. Chapter 7 of the manual contains all the approved methods for minerals in animal feeds and the references for the primary studies which established the method.

Two areas of the NIRS method require further research. First, if we are to improve the precision (accuracy can never be absolutely determined in an empirical analysis method) and, hopefully, the intended accuracy, we must improve the precision of the fiber analysis methods. Second, we must find better methods which have a molecular basis, are less empirical, and can be directly related to a physical or chemical property of the plant.

The Tecator Fibertec System is a semiautomated extraction apparatus. The word "extraction" is used here in the context of solubilization of materials from the sample by one of the following actions: Hydrolysis,

Table 4.—*Percent compositional analysis of tropical and temperate grasses*[a]

Grass (regrowth)	IVDMD[b]	Protein	Ash	NDF[c]	ADF[d]	Hemi-cellulose	Holo-cellulose	PML[e]
Coastal (4 weeks)	66.1	19.2	7.8	61.0	29.1	31.8	61.2	4.1
Coastal (8 weeks)	50.4	11.0	5.4	71.2	40.0	31.2	66.8	6.0
Coastcross-1 (4 weeks)	66.1	18.7	7.2	60.0	31.9	28.1	53.6	3.5
Coastcross-1 (8 weeks)	54.9	13.9	7.2	62.9	39.0	23.9	55.6	5.5
Bahia (4 weeks)	59.6	15.7	6.2	71.0	35.7	35.3	60.8	3.4
Bahia (8 weeks)	53.2	9.2	5.9	67.5	35.0	32.5	76.0	5.3
Pangola (4 weeks)	54.5	7.0	4.7	69.4	41.7	22.7	57.3	6.3
Pangola (8 weeks)	48.6	5.9	5.1	67.0	29.5	37.6	47.4	4.8
Average tropical	57.8	12.6	6.2	66.2	35.2	30.4	59.8	4.9
Kenhy (4 weeks)	65.6	13.2	8.3	58.2	33.6	24.7	41.6	3.2
Ken-Blue (4 weeks)	58.1	15.5	7.1	54.0	30.6	23.5	43.5	4.3
Brome (4 weeks)	64.2	14.3	8.8	56.2	34.3	21.9	51.1	4.8
Orchard (4 weeks)	62.8	14.8	8.2	57.9	33.3	24.6	43.9	4.1
Kentucky-31 (4 weeks)	62.7	14.2	8.8	58.4	31.4	25.5	45.2	3.4
Timothy (4 weeks)	66.8	13.4	8.8	55.6	34.7	20.9	42.0	4.1
Kentucky-31 (4 weeks, fall)	55.0	12.6	8.4	59.8	31.6	28.2	46.0	5.6
Kenhy (4 weeks, fall)	60.1	12.6	8.0	57.3	30.7	26.6	43.8	4.2
Orchard (4 weeks, fall)	58.8	17.8	9.7	54.0	29.4	24.6	45.5	4.8
Ken-Blue (4 weeks, fall)	61.0	17.3	6.3	57.6	7.5	30.1	40.3	4.1
Average temperate	61.5	14.6	8.2	56.9	31.7	25.1	44.3	4.3
Average standard deviation	2.62	.13	.13	1.08	1.00	1.04	1.00	.30

[a] From Barton et al. (*1976*).
[b] IVDMD = in vitro dry matter digestibility.
[c] NDF = neutral detergent fiber.
[d] ADF = acid detergent fiber.
[e] PML = permanganate lignin.

digestion, dissolution by or complex formation. The samples are introduced (weighed) into sintered glass crucibles and placed in the Fibertec hot extraction unit (up to six crucibles at a time). The preheated reagents are pumped into the crucibles and boiled in the crucible extraction column. After the allotted time, the samples are filtered and washed with no sample transfer. The Fibertec has an excellent system for washing the samples of all detergent, which eliminates one large source of error. The vacuum filtration and washing are very uniform for all samples, which helps to further reduce errors. The precision of analysis is improved overall, as shown by the SE values in table 5. The standard deviation for triplicate determinations is roughly three-fourths that of the conventional boil-and-stir procedure.

Table 5.—*Variation due to the extraction technique used for laboratory analysis of fiber*

	Conventional		Fibertec	
Constituent	**Mean**	**SE[a]**	**Mean**	**SE[a]**
NDF	68.5	0.30	64.9	0.24
ADF	39.1	.39	36.4	.34
PML	6.54	.45	5.43	.74

[a] Between rep mean square.

Any claim that calibration data have been improved can only be proven by using those data to calibrate an NIRS spectrometer. Table 6 shows the results of three different wet chemical calibration data sets on the same sample (Barton et al. unpublished results). The first data set (A) is the average of triplicate analyses by four different research laboratories (12 replicates); the second (B) is a triplicate determination by only one laboratory, the author's; the third (F) is a triplicate determination using the Fibertec in our laboratory. The precision (as indicated by the SD values) of the Fibertec (F) results was generally better than the average precision

Table 6.—*Effect of laboratory analysis data on calibration of NIRS*

Analysis	Type	Range	$\overline{SD}$	No. of wavelengths	R^{2a}	SEC[b]	Repeatability[c]
Protein	A[d]	7.21	0.62	1	0.94	1.07	0.1676
	B[e]	8.23	.18	1[e]	.95	1.05	.1434
	R[f]	7-22	.15	1	.94	1.10	.1789
NDF	A	43-76	1.99	3	.97	1.45	.2750
	B	43-75	.45	3	.96	1.76	.2693
	F[g h]	44-75	.42	3	.98	1.24	.2458
ADF	A	30-44	2.03	3	.93	1.18	.2627
	B	29-45	.25[d]	3	.85	1.84	.2994
	F	28-43	.43	3	.87	1.49	.2684
PML	A	3-12	.51	3	.92	.71	.1662
	B	2-12	.21	2	.71	1.50	.1029
	F						
ADL[i]	A	3-10	.56	3	.87	.70	.1552
	B	3-10	.11	3	.90	.67	.1201
	F	1.5-10	.30	3	.94	.57	.1368
IVDMD	A	42-76	2.57	4	.95	2.00	.6426
	B	47-69	1.04	4	.95	1.38	.5251
	R						

[a] R^2 = multiple correlation coefficient.
[b] SED = Standard error of calibration.
[c] Repeatability = measure of added signal-to-noise ratio for an additional term in the equation. If it increases with a new term, more noise than signal is being added and overfit is more of a problem.
[d] A = average of the 4 best labs (conventional method).
[e] B = conventional-method data from Athens laboratory.
[f] R = rerun of protein form Athens lab.
[g] F = Fibertec analysis data.
[h] 6 reps per sample.
[i] ADL = acid detergent lignin.

of the four laboratories and always better than that of the single laboratory determination. This is reflected in the calibration of an NIRS instrument by the R^2 and SEC values. The between-laboratory errors were large, as indicated by the average of the combined SD values from the four laboratories. The Fibertec results were obviously more precise, but the results appear to have been as good as or better than the averages from the four laboratories' data for calibration. Data in table 5 provide an internal check for within-laboratory consistency. Since crude protein was not run on the Fibertec, those values labeled "R" are repeat analyses of triplicate Kjeldahl analyses in a single laboratory at a different time and are used for internal checks of laboratory precision.

The Fibertec system offers obvious improvements in analytical precision and is easier and faster to run; moreover, the personnel needed to run the analyses is reduced. One person can run 2 Fibertec systems (12 samples per run), including weighing, and make 6 runs a day (8-hour day). In comparison, our boil-and-stir method requires 2 people during filtration and at best handled a total of 50 samples a day. The Fibertec is expensive ($15,000) and requires considerable maintenance but reduces laboratory errors and increases throughput.

Among the concerns in the second major area of research required for the NIRS method is the relationship of the compositional fraction that yields the value that we are correlating to the spectra. Recent work by Himmelsbach et al. (*1983*) has shown that solid-state ^{13}C NMR can be used to compare the relative ratios of various plant constituents between plant species. This method may be able to provide accurate data on constituents in situ. Historically, the one great limitation to using ^{13}C NMR for obtaining structural and quantitative information was that the sample had to be dissolved in a suitable solvent before the spectrum could be taken. This limitation meant that solid samples which are insoluble (for example, whole grass) were not amenable to the NMR technique. More subtly, the extraction of hemicellulose, lignin, or lignin-carbohydrate complexes (LCC) modified the in situ structure. Lignin, once isolated, was very susceptible to condensation reactions upon exposure to moisture and air. Interpretations had to be made with the knowledge that artifacts of isolation could exist. The new NMR technique of cross polarization/magic angle spinning (CP/MAS) (Schaefer and Stejskal *1979*) has made it possible to take the spectra of rigid materials. The spectra of fibrous residues and even whole grass samples can now be taken if the appropriate instrumentation is available.

Basically, there are three important constituents of plant fiber: carbohydrates, lignin, and protein. Solid-phase ^{13}C-NMR spectra of these constituents were obtained, separately or in combination, by several groups of researchers in the past 3 years (Atalla et al *1980*; Earl and VanderHart *1980, 1981*, O'Donnel et al. *1981*; Rutar et al. *1980*; Schaefer et al. *1981*; Dixon et al. *1981*) with the CP/MAS technique (Schaefer and Stejskal *1979*). Several researchers have reported that the intensities of ^{13}C-NMR signals due to protein carbonyls were proportional to values obtained for protein by standard chemical methods.

Utilizing this technique, Himmelsbach et al. (*1983*) reported the application of CP/MAS-^{13}C NMR for determining the ratio of carbohydrate, lignin, and protein between grass species. They used a single spectrum per whole plant material in the determinations. Figure 1 shows the spectra of two common forage grasses, Kentucky-31 tall fescue and Coastal bermudagrass. Both spectra contain relatively narrow peaks which correspond to the carbohydrate, lignin, and protein in the plant. The spectra in figure 2 show how the individual constituents in the whole grass can be divided into their individual components for quantitative analysis.

Since no molecular weights can be determined for the grasses and since the packing densities and a number of physical parameters cannot be easily obtained or held uniform, direct measurement of a constituent is not possible. Ratios were determined to compare the quantitative results for these two grasses. The NMR data in table 7 represent results obtained from several instruments, which were tested at two field strengths to see which gave the better results. The 50-MHz carbon data appeared to be better, and they agreed remarkably well with the NIRS data and only slightly less well with the wet chemical analyses data (CA). A prob-

Table 7.—*Ratio of fiber constituents in CBG[a] vs. KY-31[b] (CBG/KY-31) as determined by CP/MAS-^{13}C NMR, NIRS, and laboratory chemical analysis (CA)[c, d]*

Technique	Constituent		
	Carbohydrate	Lignin	Protein
CP/MAS-^{13}C NMR			
37.7 MHz	1.27	1.57	0.68
50.0 MHz	1.27	1.66	.59
NIRS[d]	1.27	1.70	.58
CA[d]	1.33	1.60	.65

[a] CBG = Coastal bermudagrass.
[b] Ky-31 = Kentucky-31 tall fescue.
[c] Each analytical value in the ratio is the result of at least triplicate replications, with standard deviation of ± 1 in percent composition.
[d] From Himmelsbach et al. (*1983*).

able explanation for this is that both spectral methods, the NIR and NMR measure responses which result in signals from the constituents in the entire plant material, whereas the wet chemical methods (which were

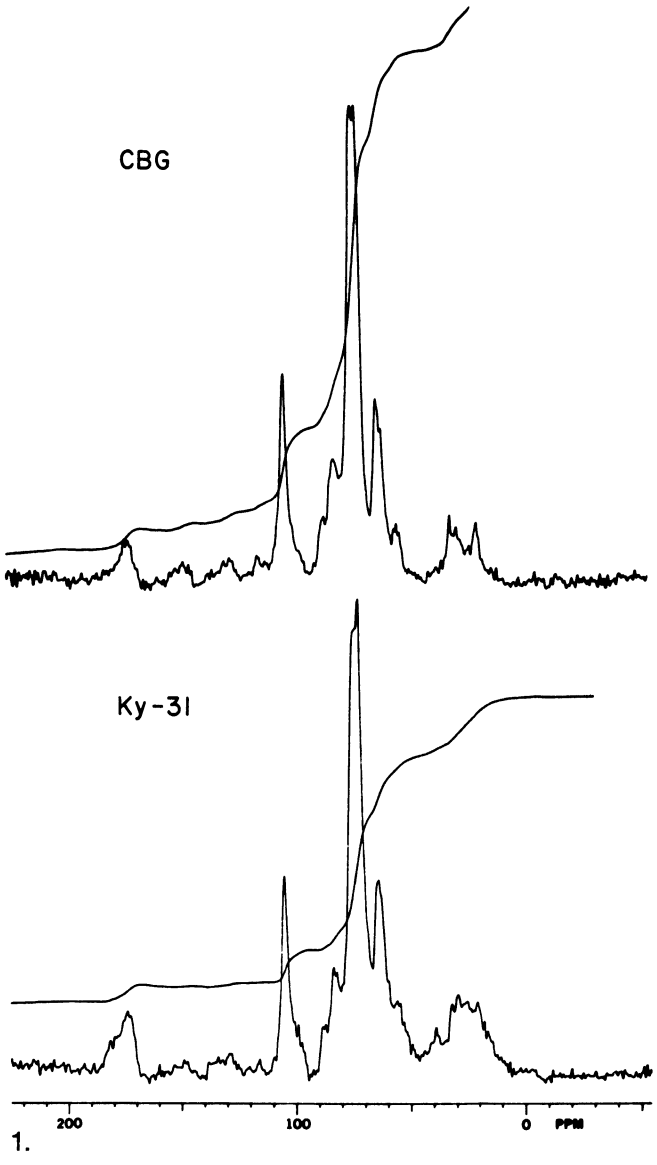


Figure 1.—Solid-state ^{13}C NMR of Coastal bermudagrass and Ky-31 tall fescue at 50 MHz.

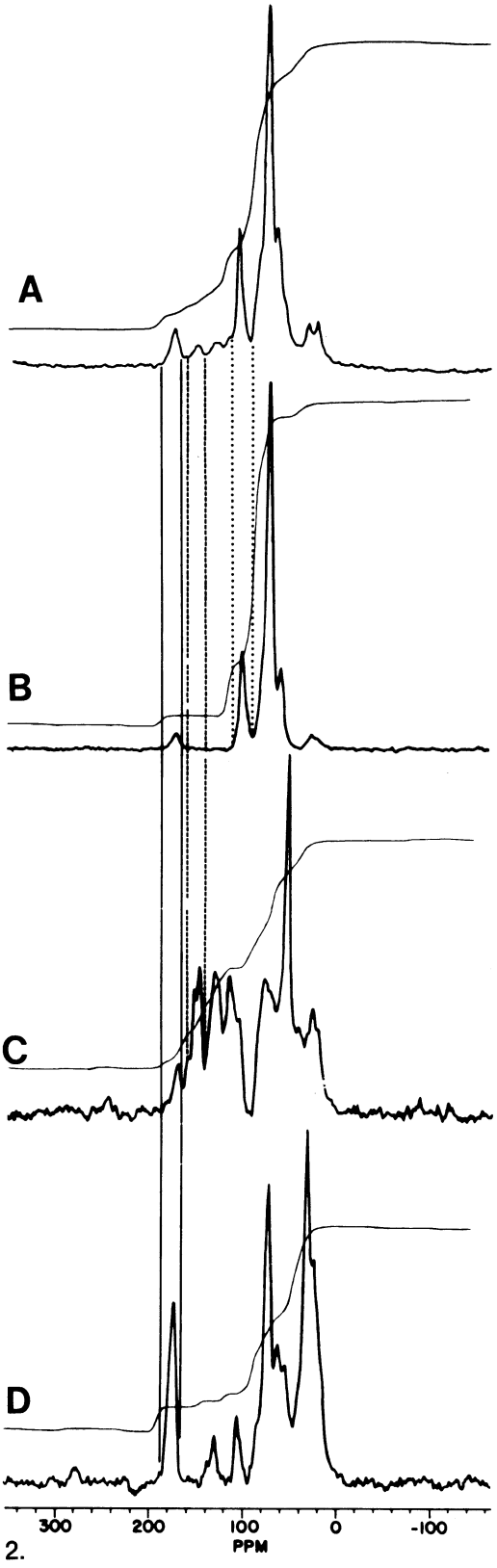


Figure 2.—Solid-state ^{13}C NMR spectrum of (*A*) Coastal bermudagrass (CBG), (*B*) CBG holocellulose, (*C*) CBG lignin, (*D*) CBG protein at 50 MHz. These spectra are used to define the regions used for quantitative determinations of carbohydrate, lignin, and protein.

empirical) measured entities which may or may not be representative of the entire plant material. In the chemical methods, as in all empirical methods, it has been assumed that all the plant materials respond to the chemical reagents identically. This, however, may not be the case. Quite possibly more low-molecular-weight carbohydrates are washed away as cell solubles from a cool-season grass species, such as Ky-31 (Kentucky-31 tall fescue), than from a warm-season grass, such as CBG (Coastal bermudagrass). Such a difference in solubility could be due to (1) structural differences in the carbohydrates of the plants and not to a true difference in carbohydrates contents of the intact plants or (2) to differences in the responses of soluble carbohydrates to the solid state technique.

Also, structural differences in lignin could also cause lignin in one species to be more susceptible to permanganate treatment. For example, in a warm-season plant like CBG, *p*-coumaryl units have been suggested to occur as side chains on polymeric lignin or even on polysaccharides as α,β-unsaturated esters (Himmelsbach and Barton *1980*). These C_9 phenolic units would be more susceptible to hydrolysis by the acid detergent reagent and removed with the hydrolyzable carbohydrates than would the true polymeric lignin. As a result there would be a slightly lower value for the CBG/Ky-31 ratio for lignin by the chemical method than there actually should be.

For protein, the instrumental techniques might produce a different ratio of protein biovalue because the standard Kjeldahl chemical method actually determines total reduced nitrogen. The Kjeldahl method could be influenced by nonprotein nitrogen in the plant, such as reduced nitrogen in free amino acids or ammonium ion in ammonium nitrate that has been absorbed into the plant after fertilization. The NIRS and CP/MAS-^{13}C-NMR methods would tend to measure protein and discriminate against free amino acids. NIRS discriminates by picking a longer wavelength for protein than amino acids. CP/MAS-^{13}C-NMR can be made to discriminate against smaller, and generally more mobile, species (Stejskal and Schaefer *1975*) and would not respond to residual ammonium nitrate.

Solid phase ^{13}C-NMR spectra can be obtained much faster than solution phase spectra for these high-molecular-weight entities (3-4 hours vs. 48 hours per sample). However, the cost of these high quality research instruments is high ($175,000 to over $300,000). Currently, CP/MAS-^{13}C NMR is the only structurally definitive (easily interpretable) method of obtaining quantitative data on plant cell walls. Another approach is to redefine the fiber of plant cell walls with an assay system based on enzymatic hydrolyses. Such an approach has been used for analyzing fiber in foods of humans (Furda *1981*, Theander and Åman *1980*). Asp et al. (*1983*) described a system for determining the soluble and insoluble fractions of dietary fiber. There is a Fibertec E system available to semiautomate the Asp procedures. It will be interesting to see how these procedures will correlate with NIRS as measures of nutrient availability.

Degradation of Plant Tissues by Rumen Micro-Organisms

While quantitative data are required for calibrating NIRS and for formulating rations, qualitative data can provide a better understanding of plant structure and intuitive knowledge of how plant cell walls are utilized by rumen micro-organisms. The primary researchers in this area are Akin and coworkers (Akin and Burdick *1973, 1981*; Akin and Barton *1983*; Akin et al. *1975, 1977, 1983*a, *1983*b, *1984*a, *1984*b). Each of their papers deals with a slightly different aspect of the degradation of forage fiber by rumen micro-organisms.

The first considerations are, What tissues are degraded? in what order? and by which organisms? In Akin and Burdick (*1975*), the percentages of tissue types and the degradation of those tissues for representative warm- and cool-season grasses are discussed. In general, phloem, mesophyll, and epidermis tissues are degraded, while the sclerenchyma and lignified vascular tissues are not. The phloem and mesophyll are degraded faster than the epidermis and parenchyma sheath, which are slowly degraded. Perhaps the most significant factor is that the cool-season grasses contain considerably more of the rapidly degraded tissues (mesophyll and phloem) and the warm-season grasses contain considerably more of the slowly degraded tissues (parenchyma sheath). There is a difference in the histochemical reaction of the lignified tissues to lignin stains (Akin and Burdick *1981*). This difference in the way lignified vascular tissue and sclerenchyma react with the acid/phloroglucinol and chlorine/sulfite reagents indicates a difference in the way the tissues are lignified. This difference is reflected in the way the tissues are attacked by rumen micro-organisms. The sclerenchyma is much less rigid and can be degraded to some extent. The tissues which are rapidly degraded can be degraded without attachment of the bacteria (that is, by extracellular enzymes). Those tissues that are slowly degraded require attachment and two types of organisms, one a cocci of the *Ruminococcus* sp. and the other a pleomorphic bacterium resembling *Bacteroides succinogenes*, which seem to be responsible for most of the attachment and degradation (Akin and Barton *1983*).

Plant cell wall digestibility depends on certain factors other than microanatomy. Maturity, plant part, and environmental conditions will affect digestibility and overall nutritive value. If a forage grass is allowed to mature and is then divided into top, middle, and bottom portions and further divided into blade, sheath, and stem, the analysis of those fractions will reveal the effect of maturity on plant quality. The work of Akin et al. (*1977*) describes the experiment outlined above with Coastal bermudagrass. In this study, with plants grown in a greenhouse for 5.5 months, there was no difference in the digestibility of blades, but the sheath and stem decreased in digestibility with increasing maturity. The chemical analyses showed that protein decreased from top to bottom and from blade to sheath to stem. Lignin content did not change for blades; but for the other plant parts, it varied inversely with protein content. Fiber content, as measured by NDF, ADF, cellulose, and hemicellulose, was not always as consistent with plant part or maturity. Later studies (Akin et al. *1983*b, *1984*a, *1984*b) focused on the range of anatomical differences within the *Panicum* genus. The genus contains species that utilize both the C_3 and C_4 photosynthetic pathways and intermediates. The results show that digestibility increased in the direction $C_3 > C_3/C_4$. Most of the generalized statements made above concerning tissue types and bacterial action against those tissues apply in the *Panicum* species studies. This genus represents a unique example in which to study many facets of the structure-composition-bacteria interactions.

The last factor to be considered is environment. Akin et al. (*1983*a) were able to study the effect of both high heat and no rain on two dry-land-tolerant species, Old World bluestem [*Bothriochloa caucasia* (Trin.) C. E. Hub.] cv. Caucasian and weeping love grass [*Eragrostis curvula* (Schrad.) Nees] cv. Morpa. When adequate moisture was provided, these two species produced abundant good quality forage; but subplots under the same temperatures with no additional moisture ceased above-ground growth, and digestibility decreased by as much as 20 percentage units for love grass. Both Old World bluestem and love grass are C_4 species and can grow in the 30° C-and-above temperatures which occurred during this study, but at these high temperatures adequate moisture was required for maximum growth and quality. These plant environmental responses to temperature and moisture affect how the plant is utilized. Such responses can best be interpreted when we understand how they are related to the chemical composition of the plant.

Plant Compositional Aspects

The microanatomical differences between cool- and warm-season (C_3 vs. C_4) grasses were examined by Akin and Burdick (*1973*). One major difference noted was that warm-season (tropical) grasses had a more developed parenchyma bundle sheath than the cool-season (temperate) grasses. In this study the sections were stained for lignin with acid/phloroglucinol reagent to identify the sites of lignification. In a subsequent study (Akin and Burdick *1975*) the percentage of each type of tissue was reported for temperate and tropical grass species. The warm-season grasses contained far more of the less easily degradable tissues than the cool-season grasses. Figure 3 shows the cross section of a typical warm-season grass, Coastal bermudagrass (CBG). Note the large and thick walled parenchyma bundle sheath on the outside of the large vascular bundle of CBG. This tissue is degraded slowly by rumen micro-organisms in CBG and, in some cases, can show a positive reaction for lignin with histochemical reagents, chlorine/sulfite in particular. The mesophyll of the warm-season grass comprises more loosely arranged cells. As shown by Akin and Burdick (*1975*), the cross-sectional area that is due to mesophyll is 50–65 percent for cool-season grasses and only 27 percent per unit for bermudagrass.

The warm-season grasses have higher fiber contents that the cool-season grasses. The values in table 4,

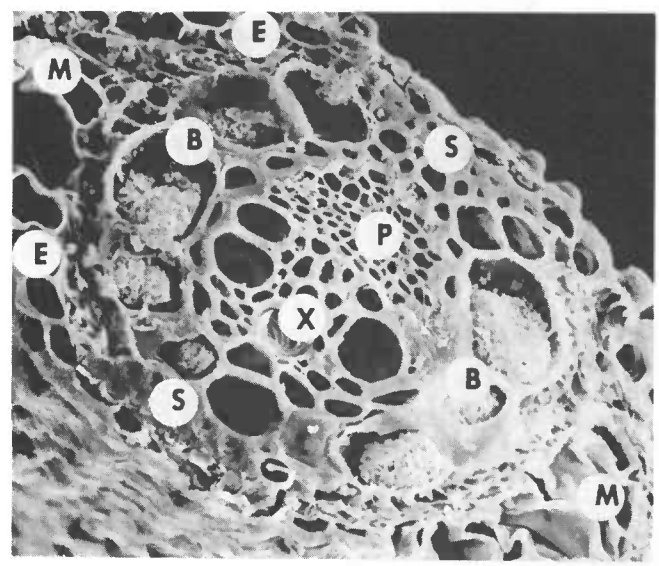


Figure 3.—Cross section of Coastal bermudagrass which shows the various tissues within the plant: sclerenchyma (S), epidermis (E), parenchyma bundle sheath (B), mesophyll (M), xylem (X), and phloem (P) × 500.

taken from Barton et al. (*1976*), reflect a 10 percentage unit average increase in NDF and 4 percentage unit increase in ADF for the warm-season grasses. These differences also occurred when the 4-week-old forages were compared. One might expect the less digestible warm-season grasses to also have a higher lignin content, and they do (4.9 vs. 3.4 percent). For the 4-week-regrowth samples, the lignin contents of the warm- and cool-season grasses were identical (4.3 percent) and the digestibilities were nearly identical (61.6 percent for warm-season vs. 61.5 percent for cool-season). Clearly, conventional compositional differences will not indicate the differences in quality or animal performance between warm-season and cool-season forages.

A possible way to resolve some of the obvious differences between digestibility and compositional analyses is to view both by the same method. Akin et al. (*1973*) showed the pattern of forage digestion by rumen micro-organisms, which is illustrated in figure 4*A* and *B*. This and subsequent studies (Akin and Burdick *1975, 1981*; Akin and Barton *1983*; Akin et al. *1975, 1977, 1983*a, *1983*b, *1984*a, *1984*b) have shown that the sclerenchyma and cuticle for both warm- and cool-season species, as well as parenchyma bundle sheath of the warm-season grass, are very slowly degraded and not fully digested within a 48-hour in vitro incubation. The microscopic (light and electron) examination of leaf sections incubated with rumen micro-organisms allows direct observation of the plant tissues that were attacked and digested. If samples are viewed after specific periods of incubation, the relative rate and mode of microbial attack can be ascertained. Microscopic examination will qualitatively describe the fiber digestion of a given species grown under the specific environmental and management conditions for a given sample but will not necessarily be related to composition. In order to relate microscopic evaluation to a particular fiber analysis, the residue after digestion by rumen micro-organisms must be determined.

Evaluation of Leaf Sections Before and After Microbial Digestion

Akin et al. (*1975*) used a scanning electron microscope to examine the tissues that comprise the residues of NDF and ADF in leaf sections from a warm-season grass (CBG) and a cool-season grass (KY-31). In these experiments, 5-mm sections of the leaf blades were treated with the boiling reagents, prepared for microscopy, and viewed with a scanning electron microscope. The percentages of tissue remaining in the leaf sections were then determined gravimetrically. For comparison, whole-leaf samples of these grasses were ground in a Wiley mill and analyzed gravimetrically for

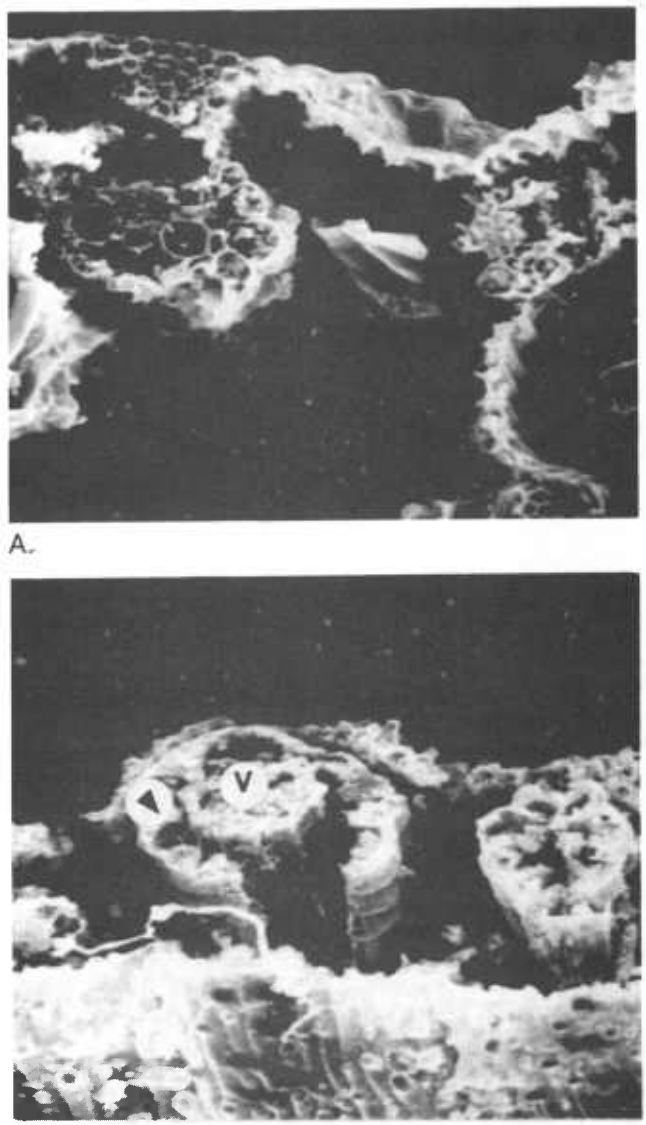
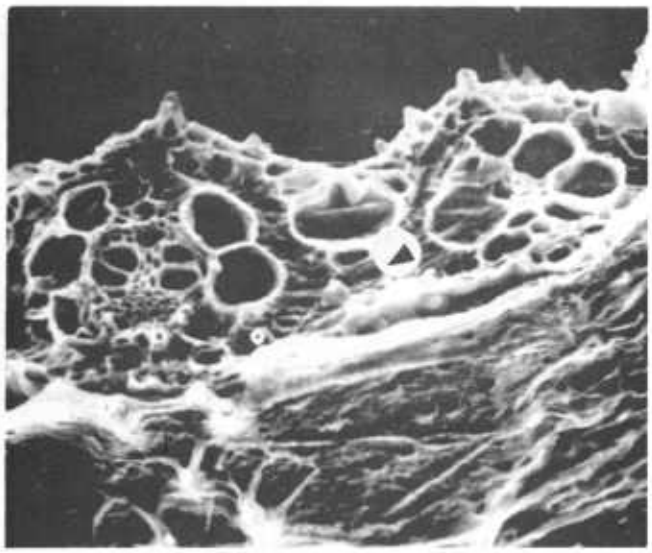
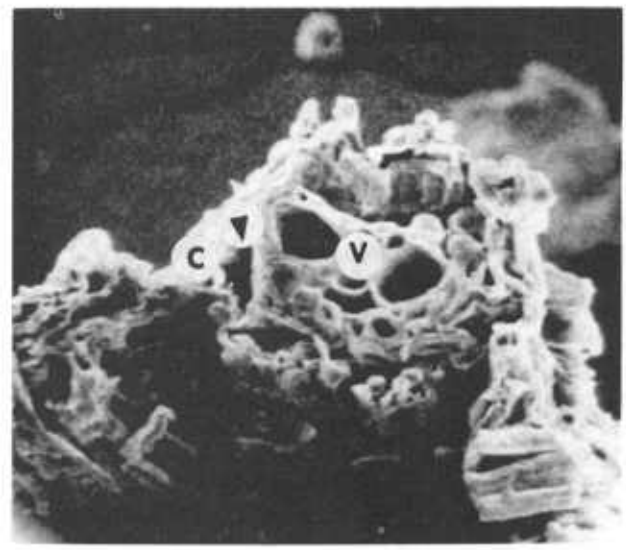


Figure 4.—*A*. Cross section of Ky-31 tall fescue incubated with rumen micro-organisms for 48 h. Lignified tissues, cuticle, and portions of bundle sheaths remain. × 150. *B*. Cross section of Coastal bermudagrass incubated with rumen micro-organisms for 48 h. The lignified vascular tissue (V) and parenchyma bundle sheath (arrow) are partially degraded. × 150.

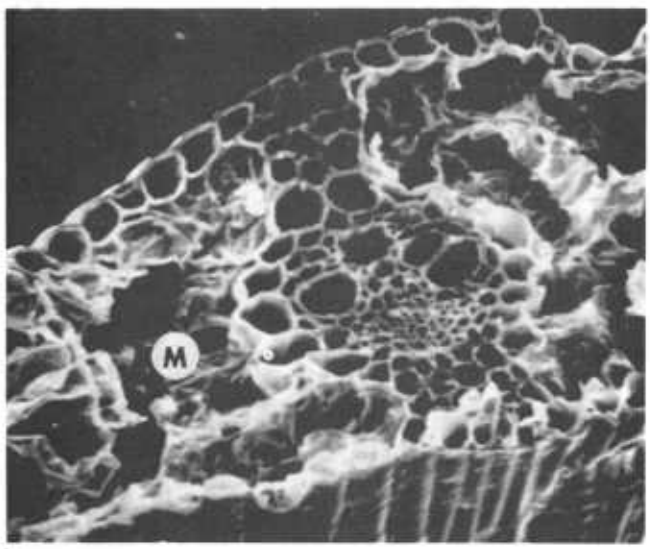
percentages of NDF and ADF. In the experiments with the intact leaf sections, the mild treatment with neutral detergent reagent left the cell walls virtually intact in CBG and slightly distorted the mesophyll in some KY-31 samples (fig. 5*A* and *C*). Much less tissue was removed from the sections than from the ground leaf blades (table 8). Apparently, therefore, the NDF reagent could



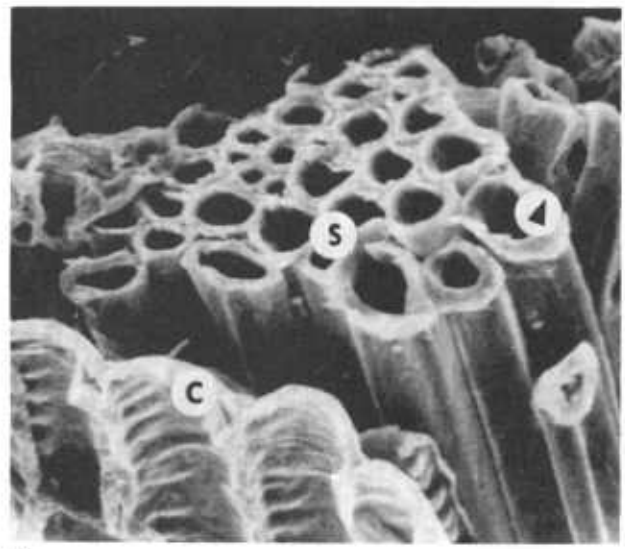
A.



B.



C.



D.

Figure 5.—*A.* Cross section of Coastal bermudagrass leaf extracted for 60 min with neutral detergent reagent. Tissues, including the mesophyll (arrow), are intact as in the control section. × 240. *B.* Cross section of Coastal bermudagrass extracted for 60 min with acid detergent reagent. Lignified vascular tissue (V) and cutinized epidermis (C) remain. Portions of the outer bundle sheath (arrow) are infrequently observed to remain. × 256. *C.* Cross section of Ky-31 tall fescue leaf extracted for 60 min with neutral detergent reagent. Tissues are similar to those in control samples; all are intact except for the mesophyll (M). × 224. *D.* Portions of tissues remaining in Ky-31 tall fescue after 60 min of extraction with acid detergent reagent. Cuticular layers of epidermis (C) and sclerenchyma (S) separated into individual cells (arrow) are seen. × 650.

not rupture the fragile cell wall membranes and could remove the cell contents only if the cell has been opened by the knife when the sections were cut. The treatment with acid detergent reagents showed some interesting differences between the species, and it contrasted with microbial digestion. For the warm-season CBG, the residue contained portions of the parenchyma bundle sheath. This tissue (fig. 5*B*), which resisted the acidic treatment, is slowly degraded by rumen micro-organisms. For KY-31, the only tissues remaining after a 60-minute treatment were cuticle, sclerenchyma patches, and pieces of vascular tissue (fig. 5*D*). This acidic digestion far exceeds the digestion of KY-31 by rumen micro-organisms. Thus, as a measure of extent of digestibility (Rohwder et al. *1978*), ADF would give too large a value for KY-31 and too low a value for CBG. Direct comparisons of quality estimated from ADF values between temperate and tropical grasses must be made with caution. The differential response of the plant cell walls to these anaytical reagents reflects differences in their availability to rumen micro-organisms; a linear response suitable for all species should not be expected.

Barton et al. (*1976*) determined correlations between digestibility and the chemical components of warm-and cool-season grasses shown in table 4. Protein and lignin contents were the most significant factors affecting digestibility. There was no correlation with NDF if both the temperate and tropical grasses were considered as one set. Clearly, specific empirical analyses do not truly measure quality when different and diverse species are belng compared.

The digestibility of ADF by rumen micro-organisms was examined by electron microscopy (Barton et al. *1981*).

Table 8.—*Percent residue of neutral detergent fiber (NDF) and acid detergent fiber (ADF) from ground whole leaves and intact 5-mm leaf sections of Coastal bermudagrass and Ky-31 tall fescue*[a]

Grass	NDF		ADF	
	Ground[b]	Intact[c]	Ground[b]	Intact[c]
Coastal bermudagrass	59.4 ± 0.3	78.3 ± 1.6	29.1 ± 0.8	25.3 ± 0.7
Ky-31 tall fescue	50.7 ± 0.6	79.1 ± 2.3	28.6 ± 0.2	27.8 ± 1.3

[a] From Akin and Barton (*1975*).
[b] Average of 12 determinations plus standard deviation for whole leaf samples ground in Wiley mill.
[c] Average of 3 determinations plus standard deviation for intact leaf sections.

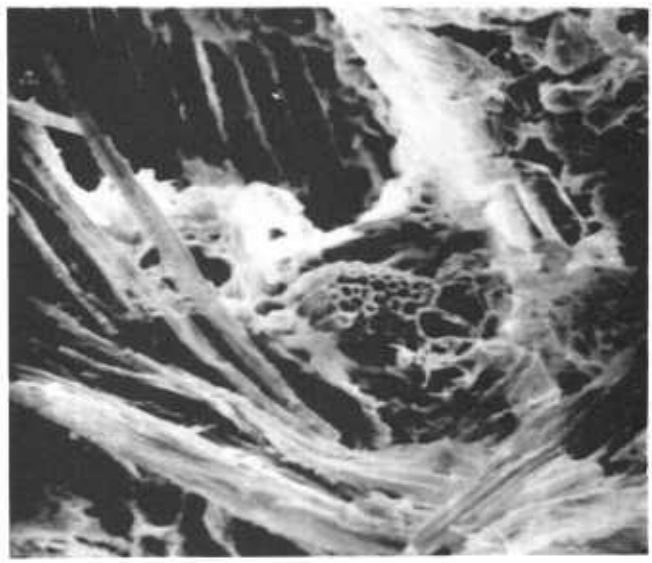


Figure 6.—Cross section of Ky-31 tall fescue extracted for 60 min with acid detergent reagent and then incubated with rumen micro-organisms for 48 h. Only cuticle, portions of vascular bundles, and sclerenchyma remain.

The patterns of degradation or removal of tissue by ADF reagents and by digestion with rumen micro-organisms were the same as those observed by Akin et al. (*1975*). In the later study, the isolated acid-detergent-treated leaf sections were subsequently incubated with rumen micro-organisms for 48 hours. The use of both treatments removed more fiber than either treatment

Table 9.—*Percentage digestibility of 5 grasses, their ADF's*[a], *and their combined dry matter disappearance (CDMD)*[b,c]

Grass	% IVDMD	% IADFD	% ADER	CDMD1	CDMD2
Coastal bermudagrass	57.41	5.66	46.56	77.24	67.08
Coastalcross-1 bermudagrass	61.17	5.91	41.17	77.16	68.53
Ky-31 tall fescue	64.54	5.76	44.73	80.40	74.93
Kenhy tall fescue	66.87	7.18	47.15	82.49	75.50
Orchardgrass	64.53	6.98	53.86	83.63	71.35

[a] Isolated ADF disappearance (IADFD) = dry matter removed from isolated ADF by rumen micro-organisms; acid detergent extracted residue (ADER) = percentage dry matter removed by ADF reagents from sample previously treated with rumen micro-organisms.
[b] CDMD1 = (NDF × ADER/100) + IVDMD; CDMD2 = (% ADF × IADFD/100) + (100 – ADF).
[c] From Barton and Akin (*1981*).

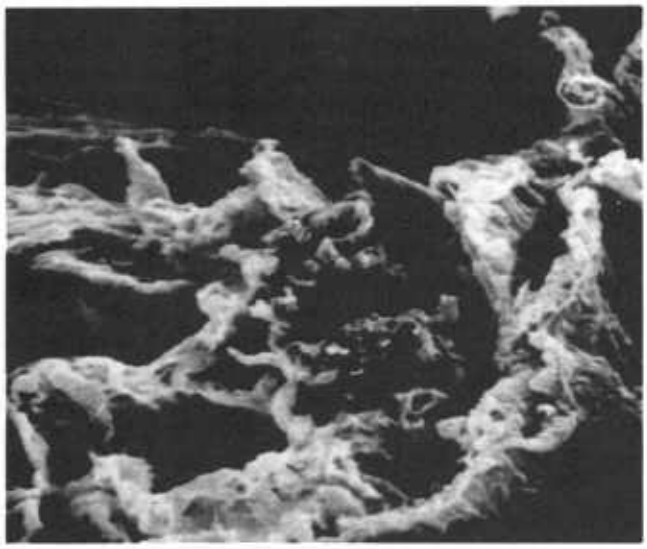


Figure 7.—Cross section of Ky-31 tall fescue incubated with rumen micro-organisms and then extracted with acid detergent reagent for 60 min. Small pieces of vascular bundles, cuticle, and some sclerenchyma remain.

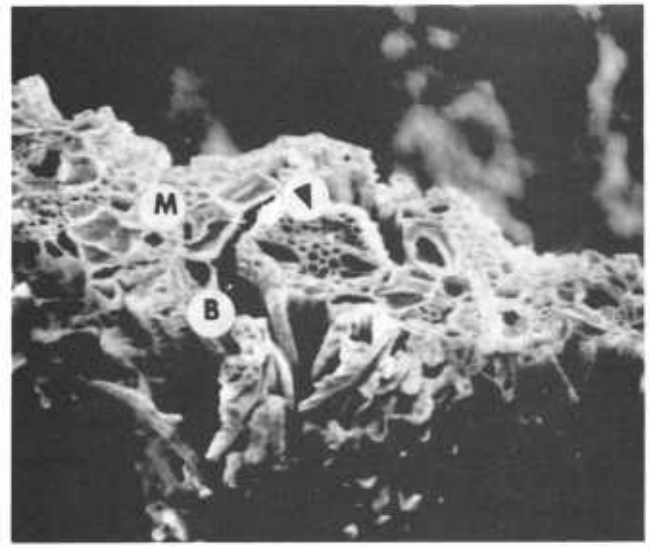


Figure 8.—Cross section of CBG-NDF treated for 30 min with buffered $KMnO_4$ reagent and demineralized for 10 min. The mesophyll (M) and phloem (arrow) are intact while the parenchyma bundle sheath (B) is almost completely separated from the rest of the vascular bundle. The sclerenchyma has been divided into individual cells and separated from the vascular bundle. × 384. (From Barton and Akin 1977.)

alone for KY-31 (fig. 6). The reverse dual treatment for KY-31 (that is, rumen micro-organism digestion followed by acid detergent) consistently removed even more tissue (fig. 7). The gravimetric results given in table 9 show a 5-11 percentage unit difference for the two treatments when the order was reversed. The qualitative data from the micrographs greatly help explain the meaning of the numerical gravimetric results. Moreover, the micrographs show that overestimation of apparent fiber digestion values can be made if no consideration is given to the effect of the dual treatment on the plant cell wall.

Traditionally, the major factors that lower forage quality and animal performance are considered to be a high amount of cell wall fiber and extent of lignification (Moore and Mott *1973*). Barton and Akin (*1977*) examined the effect of lignin removal from the cell wall on the remaining tissues and the digestion of those tissues by rumen micro-organisms. Potassium permanganate, used to oxidize lignin from ADF, was used to delignify the cell walls (NDF) of temperate and tropical grasses. This study required the authors to conduct extensive control experiments to determine exactly which of the components in the neutral permanaganate solution (such as the permanganate, buffers, or demineralizing reagents) were causing the observed effects. The results of those experiments showed that the lignified tissues were disturbed (fig. 8). The parenchyma bundle sheath was separated from the highly lignified inner

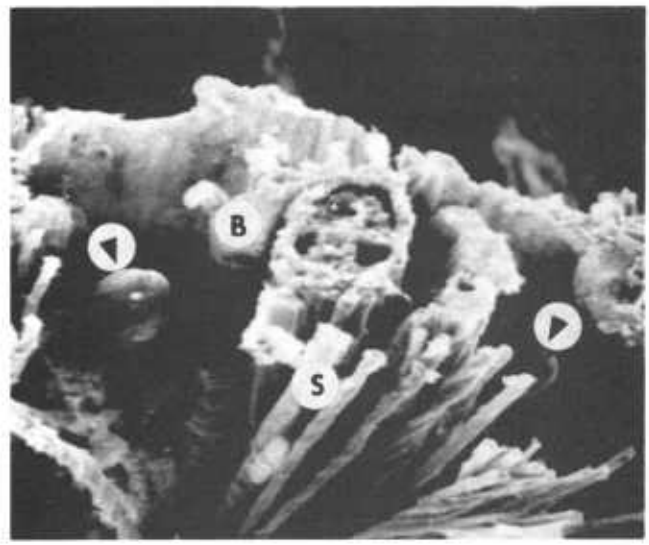


Figure 9.—Delignified CBG-NDF after 1 h of digestion. The mesophyll is completely removed, while the phloem and parenchyma bundle sheath (B) are beginning to be removed. Sclerenchyma (S) is broken apart into fiber cells and rumen protozoa (arrows) are present between vascular bundles. × 384. [From Barton and Akin 1977.]

bundle sheath; and the sclerenchyma was, in some cases, separated into individual cells. In the temperate grasses, some of the mesophyll was removed. When the permanaganate treatment was applied to NDF residue of ground forage, such that the gravimetric result of the treatment could be determined, approximately 17 percent of the material was removed. Only about a fourth of this amount was lignin; the rest of the material removed was carbohydrate. The rate and extent of digestion increased when the delignified leaf sections were incubated with rumen micro-organisms. Tissues that were usually degraded between 24 and 48 hours were extensively degraded after only 1 hour (fig. 9). Some tissues (sclerenchyma) that usually are not degraded also were digested. Thus, it is not just the amount of lignin, but the extent to which it is tied to the plant cell wall that determines the rate and extent of digestion and forage quality. These considerations are not obvious to the analyst in the usual empirical analyses, the signficance of which may be misinterpreted as a result. Methods such as those described will give analytical results that are more definitive; but the cost in time and resources would be markedly high, and the number of analyses that could be done would be quite limited.

Accuracy of Chemical Procedures

Earlier, we defined an empirical analysis as one which yielded results on a relative basis and treated all samples the same. We also stated that the result of the analysis depends on the conditions of the analysis rather than an identifiable chemical entity. Such an analysis contrasts with one in which the concentration of a real molecular entity is to be determined. In the latter case, the molecular weight of the constituent is known and its amount can be determined and expressed in units such as mmol/g, milliequivalents, mol/L, and so forth. The degree to which an analysis can be considered accurate depends on how well the procedure really describes the compositional entity. None of the current procedures adequately describes fiber. Protein can be fairly well described by a constant (5.7-6.25) times percent nitrogen, since nitrogen is a real chemical entity.

Accuracy and precision were defined earlier. We will never have a truly accurate answer for an empirical analysis; but if we carefully control the procedure, we can improve precision. The NIRS method will be very precise and accurate as long as the method used to generate calibration data is empirical. The discussion and examples cited in this handbook should be sufficient to show that NIRS is accurate enough to give compositional and animal response data that can improve the efficiency of animal and plant production.

References

Akin, D. E., H. E. Amos, F. E. Barton, II, D. Burdick.
1973 Microanotomical differences of warm-season grasses revealed by light and electron microscopy. Agron. J. 65:825.

Akin, D. E., and F. E. Barton, II.
1983 Rumen microbial attachment and degradation of plant cell walls. Fed. Proc. 42:114-121.

Akin, D. E., F. E. Barton, II, and D. Burdick.
1975 Scanning electron microscopy of Coastal bermuda and Kentucky-31 tall fescue extracted with neutral and acid detergents. J. Agric. Food Chem. 23:924.

Akin, D. E., and D. Burdick.
1973 Microanatomical differences of warm-season grasses revealed by light and electron microscopy. Agron. J. 65:533.

Akin, D. E., and D. Burdick.
1975 Percentage of tissue types in tropical and temperate grass leaf blades and degradation of tissues by rumen micro-organisms. Crop Sci. 15:661.

Akin, D. E., and D. Burdick.
1981 Relationships of different histochemical types of lignified cell walls to forage digestibility. Crop Sci. 21:577-581.

Akin, D. E., F. E. Barton, II, and S. W. Coleman.
1983a Structural factors affecting leaf degradation of Old World bluestem and weeping love grass. J. Anim. Sci. 56:1434.

Akin, D. E., J. R. Wilson, and W. R. Windham.
1983b Site and rate of tissue digestion in leaves of C_3, C_4, and C_3/C_4 intermediate *Panicum* species. Crop Sci. 23:147-155.

Akin, D. E., R. H. Brown, and L. L. Rigsby.
1984a Digestion of stem tissues in *Panicum* species. Crop Sci. 24:769-773.

Akin, D. E, L. L. Rigsby, and R. H. Brown.
1984b Ultrastructure of cell wall degradation in *Panicum* species differing in digestibility. Crop Sci. 24:156-163.

Akin, D. E., E. L. Robinson, F. .E. Barton, II, and D.
1977 Himmelsbach. Changes with maturity in anatomy, histochemistry, chemistry, and tissue digestibility of bermudagrass plant parts. J. Agric. Food Chem. 25:179–186.

AOAC.
1980 Methods of analysis. 13th ed. Association of Official Analytical Chemists, Arlington, VA.

Asp, N. G., C. G. Johansen, H. Hallmer, and M. Siljestem.
1983 Rapid enzymatic assay of insoluble and soluble dietary fiber. J. Agric. Food Chem. 31:476–482.

Atalla, R. H., J. C. Gast, D. W. Sindorf, and others.
1980 Carbon-13 NMR spectra of cellulose polymorphs. J. Am. Chem. Soc. 102:3249.

Barton, F. E., II.
1984 Advantage of the chemometric method. *In* Proceedings of the 3d Annual Users' Conference for NIR Researchers. Pacific Scientific Gardner/ Neotec Instrument Division, Silver Spring, MD.

Barton, F. E., II, and D. E. Akins.
1977 Digestibility of delignified forage cell walls. J. Agric. Food Chem. 25:1299.

Barton, F. E., II, D. E. Akin, and W. R. Windham.
1981 Scanning electron microscopy of acid detergent fiber digestion by rumen micro-organisms. J. Agric. Food Chem. 29:899.

Barton, F. E., II, H. E., Amos, D. Burdick, and R. L. Wilson.
1976 Relationship of chemical analysis to in vitro digestibility for selected tropical and temperate grasses. J. Anim. Sci. 43:504.

Dixon, W. T., J. Schaefer, M. D. Sefick, and others.
1981 Quantitative chemical composition of materials such as humic soils, lignins, and coals by high-resolution carbon-13 NMR. J. Magn. Reson. 45:173.

Earl, W. L., and D. L. VanderHart.
1980 High resolution magic angle sample spinning carbon-13 NMR of solid cellulose. J. Am Chem. Soc. 102:3251.

Earl, W. L., and D. L. VanderHart.
1981 Observation by high-resolution carbon-13 nuclear magnetic resonance of cellulose I related to morphology and crystal structure. Macromolecules 14:570.

Furda, I.
1981 Simultaneous analysis of soluble dietary fiber. *In* W. P. T. James and O. Theander, eds., The analysis of dietary fiber in food. p. 163. Marcel Dekker, New York.

Himmelsbach, D. S., and F. E. Barton, II.
1980 ^{13}C NMR of grass lignins. J. Agric. Food Chem. 28:1203.

Himmelsbach, D. S., F. E. Barton, II, and W. R. Windham.
1983 Comparisons of carbohydrate, lignin, and protein ratios between grass species by cross polarization-magic angle spinning carbon-13 nuclear magnetic resonance. J. Agric. Food Chem. 31:401–404.

Kowalski, B. R. ed.
1977 Chemometrics: theory and applications. ACS Symposium Series No. 52, American Chemical Society, Washington, DC.

Marten, H., and H. Russwurm, Jr. eds.
1983 Food research and data analysis, Applied Science Publishers, New York.

Mohsenin, N. N.
1984 Electromagnetic radiation on properties of foods and agricultural products. Gordon and Breach Science Publishers, Inc. New York.

Moore, J. E., and G. O. Mott.
1973 Structural inhibitors of quality in tropical grasses in anti-quality components of forages. *In* A. G. Matches, ed., Anti-Quality Components of Forages, p. 53–98. Crop Science Society of America, Madison, WI.

Norris, K. H.
1983a Extracting information from spectrophotometric curves. Predicting chemical composition from visible and near-infrared spectra. *In* H. Martens, H. Russwurm, Jr. eds., Food Research and data analysis, p. 95–114. Applied Science Publishers, New York.

Norris, K. H.
1983b Multivariate analysis of raw materials. *In* L. W. Shemilt, ed., Chemistry and world food supplies: the new frontiers, CHEMRAWN II [Chemical Research Applied to World Needs]. p. 527–535. Pergamon Press, New York.

O'Donnel, D. J., J. J. Ackerman, and G. E. Maciel.
1981 Comparative study of whole seed protein and starch content via crosspolarization-magic angle spinning carbon-13 (isotope) nuclear magnetic resonance spectroscopy. J. Agric. Food Chem. 29:514.

Rohweder, D. A., R. F. Barnes, and N. Jorgensen.
1978 Proposed hay grading standards based on laboratory analyses for evaluating quality. J. Anim. Sci. 47:747-759.

Rutar, V., R. Blinc, and L. Ehrenberg.
1980 Protein content determination in solid organic materials by proton-enhanced magic angle sample spinning carbon-13 NMR. J. Magn. Reson. 40:225.

Schaefer, J., M. D. Sefcik, E. O. Stejskal, and others.
1981 Macromolecules 14:557.

Schaefer, J., and E. O. Stekskal.
1979 High-resolution ^{13}C NMR of solid polymers. *In* G. C. Levy, ed., Topics in carbon13 NMR spectroscopy, vol. 3, p. 283-324. Wiley-Interscience, New York.

Stejskal, E. O., and J. Schaefer.
1975 Removal of artifacts from cross-polarization NMR experiments. J. Magn. Reson. 18:560.

Theander, O., and Åman, P.
1980 Chemical composition of some forages and various residues from feeding value determination. J. Sci. Food Agric. 31:31-37.

Van Soest, P. J.
1963a Use of detergents in the analysis of fibrous feeds. I. Preparation of fiber residues of low nitrogen content. J. Assoc. Off. Anal. Chem. 46:825.

Van Soest, P. J.
1963b Uses of detergents in the analysis of fibrous feeds. II. A rapid method for the determination of fiber and lignin. J. Assoc. Off. Agric. Chem. 46:829.

Van Soest, P. J.
1967 Development of a comprehensive system of feed analyses and its application to forages. Anim. Sci. 26:119.

Van Soest, P. J.
1973 The uniformity and nutritive availability of cellulose. Fed. Proc. 32:1804.

Van Soest, P. J., and R. H. Wine.
1967 Use of detergents in the analysis of fibrous feeds. IV. Determination of plant cell-wall constituents. J. Assoc. Off. Agric. Chem. 50:50.

Van Soest, P. J. and R. H. Wine.
1968 Determination of lignin and cellulose in acid detergent fiber with permanganate. J. Assoc. Off. Agric. Chem. 51:780.

3. In Vivo and In Vitro Measurements of Forage Quality

S. W. Coleman and W. R. Windham

Components of Forage Quality

The feeding value of a forage or other feedstuff is the product of intake, digestibility, and utilization (Raymond *1969*). The common forages fed to ruminants have been evaluated much more extensively for energy content, digestibility, and even utilization than for intake (Waldo *1969*). Heany (*1969*) suggested that combining digestibility and intake into a single index provides a means of evaluating the feeding value of forages more effectively than of evaluating either alone.

When using digestible energy intake (DEI) as an index of forage quality, the relative contributions of intake and digestibility are not the same. Milford and Minson (*1965*) found that the digestible dry matter intake (DDMI) of tropical grasses was more correlated with intake of dry matter than its digestibility. Crampton et al. (*1960*) reported that variations in intake accounted for 70 percent of the variability in the nutritive value index. Crampton (*1957*) and Ventura et al. (*1975*) agreed that intake is the more important factor in determining quality, but intake of a given forage is more variable between animals than is digestibility (Blaxter et al. *1961*, Minson et al. *1964*, Heaney et al. *1968*). The importance of voluntary intake, however, does not imply that digestibility is not important in determining DEI or other expressions of quality. Blaxter et al. (*1961*) calculated that under ad libitum feeding conditions, a change in digestibility of DM from 50 to 55 percent resulted in a 100-percent increase in weight gain.

Much work has been done attempting to relate chemical composition of forages to forage quality. Low protein content has been considered the limiting factor in controlling feed intake (Milford and Minson *1965*). However, this generally occurs only when crude protein (CP) content of the forage falls below 6–7 percent of the DM (Minson and Milford *1967*). Above this critical level, rumen fill is considered the primary determinant of intake in ruminants (Campling et al. *1961*, Conrad *1966*), especially with lower quality forages. When digestible energy intake satisifies the requirement for maximum production, control of intake is no longer based on rumen capacity but on chemostatic or thermostatic factors (Montgomery and Baumgardt *1965*, Blaxter et al. *1961*).

When rumen distention is the mechanism of intake control, the structural components of forage, which define digesta volume, are considered to be the cause of the distention. The systems of fibrous feed analysis devised by Van Soest and associates have become accepted procedures for evaluating forage quality. These analyses attempt to divide feeds into digestible and indigestible fractions. Basic to the problem of using cell wall constituents (CWC) to predict forage quality, however, is the fact that CWC, or any other representative of the fiber portion, is not a nutritionally uniform fraction as defined by Lucas et al. *1961*. In general, the digestible fraction is composed of a soluble, readily available portion, and an insoluble, partially available portion. The soluble fraction is usually referred to as the cellular contents and is soluble in neutral detergent (Goering and Van Soest *1970*). The insoluble fraction contains substances of the plant cell walls that may or may not be digested by rumen micro-organisms and is comparable to the neutral detergent fiber fraction (NDF). The most highly lignified and least digestible fraction is acid detergent fiber (ADF), which resides in the plant secondary cell wall. The unavailability of this fraction is perhaps due to physical-chemical factors such as encrustation, lignification, crystallinity of cellulose, and the organizational structure of the forage cell wall fraction.

Laboratory Prediction of Forage Quality

The cell wall fraction constitutes the structural part of the plant and is the least and most slowly digestible portion. Thus, it determines the space-occupying capacity of a forage or feed (Van Soest *1965*) and should afford the best predictor of intake. Van Soest reported results from 82 forages (6 plant species) in which intake was correlated with various chemical components. Total correlations over all species showed CWC to be best related ($r = 0.65$, $P > 0.01$) to intake. Correlations with intake of all components (lignin, ADF, protein, and cellulose) were similar within species, indicating the uniform influence of maturity on forage quality. However, correlations across species were more variable. Regression analyses indicated that the relationship between intake and CWC was curvilinear, with the influence of CWC being markedly depressed when it constituted less than 50 percent of the DM. This suggests that CWC, representing the total fibrous part of the forage, limited intake when the proportion of these constituents increased to more than 55–60 percent of the dry matter. These relationships are consistent with observations regarding the existence of a point in the intake-fiber mass relationship where fiber mass ceases to affect intake (Conrad et al. *1964*, Montgomery and Baumgardt *1965*).

Table 10.—*Relationship of laboratory techniques to digestibility of forage samples*

Method	N	R^2	Std. error estimate	References
Acid detergent fiber	122	0.22		Van Soest et al. *1978.*
Acid detergent fiber	30	.07	5.6	Laredo and Minson *1973.*
Neutral detergent fiber	30	.06	5.5	Laredo and Minson *1973.*
Lignin	30	.20	5.1	Laredo and Minson *1973.*
Lignin[a]	15	.81	2.3	Laredo and Minson *1973.*
IVDMD[b]	36	.92	2.2	McLeod and Minson *1974.*
NIRS	76	.95	2.5	Norris et al. *1976.*

[a] Stem only.
[b] In vitro dry matter digestibility (Tilley and Terry *1963*).

Recognition of the importance of digestibility led to development of the two-stage in vitro technique (Tilley and Terry *1963*) and subsequent modifications (Monson et al. *1969*, Marten and Barnes *1979*). By applying this technique, researchers have been able to conduct systematic studies of factors influencing digestibility of forages, such as variation between species and varieties, and to estimate genotypic variation and heritability. The improved precision and acceptability of the technique over prediction from chemical analyses (table 10) added a whole new realm of selection criteria in plant breeding programs.

Reid (*1966*) reviewed the state of the art of forage evaluation. At that time, in vitro fermentation procedures were gaining acceptance for estimating digestibility, and Van Soest had completed his series of articles describing chemical fractionation of feeds. No laboratory technique was available for adequately estimating intake. To date, the state of the art is approximately the same. A technique for screening feeds and forages for components important to animal production is needed. The technique should be fast, routine, require very small amounts of sample, and precisely predict the parameter of interest.

Though the in vitro technique has been widely received and used for estimating digestibility, little has been accomplished to incorporate an estimate of intake into forage quality assessment. One of the problems of assessing intake is the inherent animal variability and bias due to class and status of the animal. Intake assessment could very well lead to greater improvement in the quality of warm-season forages than has digestibility assessment due to their high fiber, long ruminal residence time, and slow rate of digestion. Rate of digestion and rate of passage are important factors relative to mechanisms which control intake (Waldo et al. *1972*), but they are not causative agents. Several efforts have been made to identify or characterize the causative agents. It is likely that, if identified and quantified, these causative agents may be particularly sensitive to NIRS analysis. Balch (*1971*), Sudweeks et al. (*1975*), and Welch and Smith (*1969*) suggested that rumination time or time spent chewing is related to fibrousness or coarseness of roughages. Welch and Smith found a significant correlation ($r = 0.99$) between duration (minutes) of rumination and CWC intake. These results suggest that some factor other than presently known chemical fractions influences rumen fill, which in turn influences intake. Lignin content per se probably has little effect on extent of digestibility, but the amount of lignified (encrusted) tissue was implicated as being very important on the basis of microscopic evaluations (Akin et al. *1974*, De la Torrie et al. *1974*). The physical characteristics of forage plants which influence the rate and extent of digestion, particle size reduction, and rumen clearance rate must be understood if intake is to be predicted by inexpensive laboratory methods. A few potential techniques for predicting intake are listed in table 11. Though NIRS is listed, few data have been published relating intake to near infrared spectra, due primarily to the difficulty in obtaining sufficient samples of known intake. One might expect interactions of NIRS prediction to forage type similar to those shown by Moore (*1977*), who predicted intake and digestibility of tropical forages from chemical components by using equations for temperate forages. He showed that equations to predict the digestibility of hay from ADF grossly underestimated the digestibilities of tropical hays. We have had greater difficulty calibrating NIRS for chemical composition of warm-season forages than was observed for cool-season forages.

Table 11.—*Relationship of forage intake to laboratory analytical techniques*

Method	N	Intake, g/$W^{0.75}$ Species	Mean	SD[a]	R^2	Std. error of estimate	References
Artificial mastication	14	Sheep	49.5	4.3	0.88	1.9	Troelson and Bigsby *1964.*
Grinding energy	30	Sheep	48.7		.63	8.2	Laredo and Minson *1973.*
Bulk density	30	Sheep	48.7		.49	9.7	Laredo and Minson *1973.*
Acid detergent fiber	12	Sheep	51.7	6.9	.66	5.2	Lippke *1980.*
Neutral detergent fiber	12	Sheep	51.7	6.9	.09		Lippke *1980.*
NIRS	21	Cattle[b]	52.6 -112.3[c]		.72	9.6	Ward et al. *1982.*
NIRS	76	Sheep	39.7 -114.3[c]		.62	7.8	Norris et al. *1976.*

[a] SD = Standard deviation of mean of animal data.
[b] Grazing with total fecal collection.
[c] Range of the data.

Nonforage Factors Affecting Forage Quality Measurements

Quality estimates of a feed obtained by bioassay, especially those using animals, may include factors external to the feed itself and thus may not be included in the NIRS spectral information. Therefore, these factors contribute to imprecision and potential bias when NIRS, or any chemical analysis, is used to predict animal measurements, since NIRS can detect only what is contained in the forage or feed sample. There are many such variables, usually a part of the conditions of the experiment in which the parameters are estimated. Some of the more notable ones include animal, environment, levels of feeding, and feed processing.

Animal Factors

Both intake and digestibility of a feed may be influenced by animal factors. Though intake is probably a larger factor in digestible-energy intake (Crampton *1957*, Heaney et al. *1968*), intake of a given forage is more variable among animals than is digestibility. This variation among animals may be due to one or more of the following factors: Animal weight (Heaney *1969*), fatness (Bines et al. *1969*), physiological rumen volume (Purser and Moir *1966*), and retention time of organic matter in the rumen (Campling et al. *1961*, Hungate *1966*).

The influence of animal weight can be reduced by simply expressing the absolute daily intake in terms of body weight. Much controversy has occurred over the years concerning the exponent of weight that should be used to form animal "equivalents." Brody (*1945*) suggested that weight raised by the 0.73 power was sufficient for all mammals from mice to elephants. However, Kleiber (*1961*) suggested that weight be raised to the 0.75 power for easier calculation. Both of these researchers were referring primarily to fasting metabolism as the measurement of interest; and in Brody's curve, cattle and sheep were the largest outliers. Others have suggested that intake is related more to mass than to surface area and that, therefore, a simple division by weight is sufficient (Conrad et al. *1964*, Taylor and Young *1968*). Karue et al. (*1973*) fed cattle a variety of feeds from high concentrate to high forage and found that for forages, more animal variation was accounted for as the exponent approached unity.

When sheep or other small ruminants are used as pilot animals for cattle, other factors may become important. Most investigators believe that sheep and cattle differ significantly in intake (table 12). In general, sheep tend to be more senstive to physical characteristics, plant parts, and so forth, and are more likely to have lower intakes of poor quality forages, especially poor quality silages. Heaney (*1979*) suggested that cattle tend to digest roughages more efficiently than do sheep and that this difference is accentuated as the digestibility of the forage decreases. However, he suggested that relative differences in digestibility between herbages are reasonably consistent regardless of whether they are determined with sheep or cattle so long as digestibility is above 45 percent. The data in table 12 suggest that in some cases there may be serious problems extrapolating sheep data to cattle. There does appear to be adequate agreement between sheep and cattle for ranking intake of the grasses and legumes usually used for pasture and conserved forage (Heaney *1979*). Relative differences in feed value between herbages should be similar, but absolute values may differ (Heaney *1979*). These observations are important for NIRS, for it suggests that experiments with sheep and those with cattle cannot be combined, at least without some scaling factors which are unknown at the present time.

Table 12.—*Comparison of sheep and cattle for in vivo measurement of digestibility and intake*

Feed	Intake (%BW)		Digestibility[a]		Reference
	Sheep	Cattle	Sheep	Cattle	
Dehydrated alfalfa[b]	116	129	59.8	51.7	Anderson et al. *1977.*
Sedge[b]	54	100	51.3	59.8	Anderson et al. *1977.*
Corn silage	1.65	1.67	67.3	68.9	Colovos et al. *1970.*
Corn silage	1.65	2.02	65.2	67.2	Colovos et al. *1970.*
Corn silage	1.74	1.90	72.5	65.2	Colovos et al. *1970.*
Corn silage	1.52	1.72	71.1	66.0	Colovos et al. *1970.*
Alfalfa[c]	2.51	3.31	67.1	51.6	Kilmer et al. *1979.*
Grass hay	2.51	2.72	65.3	59.5	Kilmer et al. *1979.*

[a] Dry matter.
[b] Intake is $g/W^{0.75}$.
[c] Energy digestibility.

The physiological status of the animals also may influence intake and digestibility measurements (Foot *1972*, Forbes *1977*). Fatness, pregnancy, lactation, and stage of growth all certainly influence intake. The mechanisms include energy demand and rumen capacity. Digestibility is less affected by these physiological factors, although the amount of omental fat has been suggested to reduce digestibility in very fat animals. Though the magnitude of the effects of phsiological status on feed intake and digestibility may be less than those of species or size, they may also be harder to adjust for since the effects, singly and combined, have not yet been well defined.

Rumen capacity may influence intake. This is one characteristic that distinguishes dairy breeds from beef breeds, the dairy breeds having a larger rumen capacity per kilogram of body mass. Digestibility may only be affected if the larger capacity is combined with a lower relative intake so that residence time is longer. Rate of passage may be considered as both a feed factor and an animal factor; animal genetic differences in the ability to reduce particle size and remove undigested residues from the rumen may exist.

Environmental Factors

Many environmental factors influence quality measurements with animals. The environment includes all nongenetic or nonanimal factors. The physical surroundings and climatic conditions during an experiment to determine digestible nutrient intake may influence the measurements. It is known that voluntary intake by animals can vary markedly with season (Forbes et al. *1979*, Evans and Potter *1984*) and that the requirements or demand for feed increases with decreasing temperature. This variability presents a particular problem in the calibration of NIRS for intake and digestibility since a large number of forages and feeds usually cannot be fed in a single balanced experiment. A long period including many feeding experiments may be required. Evans and Potter evaluated a procedure to minimize these effects by the inclusion of a control feed. For NIRS calibration, however, it would be desirable to use samples previously evaluated in trials which had no control feed. Therefore, an alternate method of reconciling seasonal and other environmental effects would be desirable. None are known at this time, but inclusion of discrete factors such as year, trial, and season in the calibration matrix can identify either offsets or interaction (Minson et al. *1984*). The practical use of the myriad of factors as discrete effects may be impossible.

Level of Feeding

The level at which a feed, especially a forage, is fed may significantly affect its digestibility. Of course, intake measurements must be made under ad libitum conditions. As the level of intake increases, the rate of passage is increased and, generally, digestibility is decreased (Blaxter et al. *1956*, Coleman et al. *1978*). Extent of digestion is a function of rate of digestion and retention time. As intake increases, retention time decreases; thus, time allowed for digestion of a particle and, hence, digestibility are reduced. If digestibility and intake are to be combined in an index for energy intake, then digestibility should be obtained at ad libitum intake. However, if digestibility is to be used as the sole criterion of forage quality, more uniformity of the estimates is likely if all the forages are fed at a given level. The level is usually expressed as a multiple of the maintenance level for the animal which is being fed. Measurements of digestibility made in this way are more dependent on the fermentability of the forage by rumen microbes than on the combination of factors, both plant and animal, including rumen fill, passage,

and residence time. In vitro estimates generally mimic very well measurements made at restricted intake.

Feed Processing

The kind and degree of physical (and/or chemical) processing of a feed may influence measurements of intake and digestibility. Much attention has been given to the effects of physical form such as grinding and pelleting on nutritive value and animal performance. Excellent reviews have been published by Putnam and Davis (*1961*), Beardsley (*1964*), and Moore (*1964*).

An understanding of those physical characteristics of forage plants which influence rate and extent of digestion, particle size reduction, and rumen clearance rate is necessary in order to be able to predict intake using inexpensive laboratory methods.

Physical processing such as pelleting, flaking, steaming, and so forth, pose special problems to prediction of intake by NIRS. Generally, these procedures may not alter the near-infrared-sensitive chemistry of the plant, and their effects are thus transparent to NIRS prediction of quality. However, particle size after laboratory mill grinding is influenced by these processing methods as is total reflected energy. Large differences in intake and digestibility by animals usually occur due to processing and will cause serious biases in prediction of the parameters by NIRS. Except for special industries, such as those producing dehydrated alfalfa and industrial byproducts, most industries seldom use these processing methods due to the high energy inputs.

In Vivo Intake and Digestibility Trials

Intake, digestibility, and dry matter turnover could be calibrated on NIRS instruments by using a variety of forages in which quality is variable. Variation in quality may be generated by using different plant species, maturity, methods of preservation, and so forth. These forages must then be fed in such a way as to minimize both random and nonrandom variations caused by the factors previously discussed. An indepth discussion of procedures for determining digestibility is available (Schneider and Flatt *1975*). However, little has been reported on the conduct of intake trials.

Equipment

Accurate feeding of weighed amounts of feed and, in the case of digestion trials, accurate recovery of the excreta are important in in vivo experiments. Points to be considered have been mentioned in the previous discussion. Some important ones to be considered in planning experiments are (1) feeding the experimental animals, (2) allowing access to drinking water, (3) collecting uncontamined feces, (4) comfort of animal, (5) equipment needed, (6) collecting and preserving experimental feeds, and (7) minimizing nonfeed factors.

Equipment used for measuring voluntary intake may be very simple, though the conduct of such trials may be quite laborious. A restraining stall which will hold the required number of animals with a feed bunk and water will suffice. Minson and Cowper (*1977*) described an elaborate mechanism for measuring voluntary intake in cattle by automatically removing the refused feed hourly. The use of electronic headgates to allow only 1 animal from a pen of 12 to eat from an individual feed container have also been used. However, when coarse forages such as long hay or silage are fed, large feed bunks are necessary to prevent wastage.

Equipment for digestion trials is a bit more sophisticated, though it may still be simple. Since digestibility of a feed is defined as the difference between what is fed and what is excreted in the feces, some mechanism must be provided for collection and quantification of the feces. Generally, stalls which allow collection of the urine and feces in separate containers are used.

Animals

Regardless of whether the test animal is to be a model, such as sheep, or the end recipient of the feed, such as cattle, certain precautions must be observed. If the trials are to be conducted over a long period, the test animals must be as uniform as possible with respect to breed, size, type (that is, dairy or beef), long or short wool or hair, fatness, maturity, and condition (pregnant, lactating, growing). The animals should be obtained prior to initiation of the trial and adapted to as many of the experimental conditions as possible. Large cattle may be difficult to adapt to elevated digestion stalls and may exhibit reduced intake, weight loss, and other problems.

Feed Processing

Usually, NIRS spectra will not be influenced by the methods used to process the feeds. However, as earlier discussed, processing often has a considerable effect on intake and digestibility and may cause bias when NIRS is used to predict quality of processed feeds. Therefore, standardized procedures should be used. For all kinds of forages to be used in a single calibration, that is hays, silages, and fresh forages, similar particle size should be offered. Most silage harvesters have variable length choppers, and typical silage is chopped to 1-5 cm. Both hay and fresh forage can also be chopped; therefore, it is recommended that a theortical

cut of 2.5 cm be used. The typical particle size of haylage and silage is larger than the theoretical length of cut, whereas particle size of dry hay will tend to be less than the theoretical cut due to shatter. However, these procedures will go far in reducing the effects of particle size.

Grinding, pelleting, flaking, and/or steaming are all special processes. They change the fiber structure, increase rate of passage, decrease extent of digestion, and sometimes alter protein to undegradable forms. Therefore, except for special circumstances where an entire data set is to built from a processed feed, these processes should not be included.

Level of Feeding

Level of feeding may influence rate of passage and extent of digestion in a manner that is independent of those factors that can be determined by NIRS. Therefore, calibration of the NIRS may be difficult. Separate intake and digestion trials partially overcome the problem, but theoretical questions still arise. The dynamic interactions among level of intake, rate of passage, and rate of digestion pose special problems for calibration of NIRS. Consequently, it may be preferable to calibrate for digestion and passage rate coefficients rather than to calibrate for extent of digestion. The extent of digestion to be expected for given levels of intake could then be calculated, and therefore the calibration would be adequate at all levels of intake. However, current techniques for estimating rates of passage and digestion are too imprecise and inaccurate for general use.

Number of Animals Required

Animal trials are time consuming and expensive. Further, as discussed above, within animal variation can be substantial. Few quantitative data are available for variation in intake. The standard deviation of a dry matter digestion coefficient is usually between 1.0 and 1.3 percentage units (Forbes et al. *1946*, Raymond et al. *1953*). Using three sheep per digestibility measurement will give 80 percent chance of detecting digestibility differences of 3-4 percent at the 5 percent level of significance. Variability of coefficients obtained with cattle may be slightly higher.

Intake measurements are less precise (Heaney et al. *1968*), with standard deviations of 7 to 10 intake units ($g/W^{0.75}kg/day$). Therefore, about 12 sheep are required to detect differences of 10 intake units. Usually three sheep are required for digestibility trials and six sheep if intake is to be measured (Heaney *1979*). Variation with cattle may be slightly more, but we have successfully used six steers for intake, provided a period of 60 to 100 days is used for the estimate. This length of time is also marginally sufficient for rate of gain.

Procedures for Cattle Trials

Intake. Intake is measured by providing the test feed to animals for a specified duration. Sixty plus days for cattle and 30 plus for sheep are recommended. Excess feed (10-20 percent) is provided to assure that the animal is truly eating ad libitium.

At least six animals each weighing from 250-350 kg will be individually fed each hay for 60 days ad libitum to determine voluntary intake. Forage can be offered initially at 4 percent of body weight and the amount reduced daily by 1 kg until the refusal is approximately 15 percent each day. Thereafter, refusal in excess of 15 percent will be required for 2 consecutive days to reduce feed on offer. Immediate adjustment upward will be made if refusal is less than 15 percent.

Digestibility. At least four animals will be fed each hay. Animals will be allowed 9 days to adjust to rations to determine ad libitum intake for the conditions. Thereafter, forage offered will be constant at the level determined to be ad libitum during the preliminary period. Beginning 2 days after the end of the preliminary period, feces and urine will be collected daily for 7 days. Five percent aliquots will be collected and refrigerated for later analysis. Samples of feeds refused and feces will be analyzed for dry matter, ash, and crude protein according to AOAC (*1970*) procedures. Acid detergent fiber, neutral detergent fiber, lignin, and cellulose will be determined according to Goering and Van Soest (*1970*).

Rate of passage. Rate of passage can be estimated during a normal digestion trial by taking extra fecal grab samples. Since it is closely involved with intake and digestion, its measurement and calibration on NIRS seems highly recommended. On day 16, each animal will be dosed with 8-10 g ytterbium chloride ($YbCl_3 . \times H_20$) attached to 500 g of the forage being fed. This material will be given at the evening feeding and will be flavored with an ingredient such as molasses to induce eating. After 12 hours, fecal grab samples will be obtained every 4-6 hours for 2 days and every 8-12 hours for the remainder of the collection period.

Rate of digestion. This parameter may also have value in explaining discrepancies in calibration and prediction of intake and digestion by NIRS. In vitro digestibility rate will be determined on each hay by in-

cubating it in rumen fluid for 0, 3, 6, 12, 18, 24, 36, 48, 72, and 96 hours and weighing the residual NDF.

In Vitro Digestibility Estimates

In vitro rumen fermentation methods for forage quality evaluation can be used to quite rapidly predict in vivo digestibility, and they require much smaller samples than those needed for large-animal in vivo studies. A number of reviews have been published which describe the development, modification, and application of in vitro rumen fermentation techniques (Barnes *1965*, Homb *1963*, Johnson *1963*, Pigden *1969*, Raymond *1969*, Shelton and Reid *1960*). Barnes (*1973*) and Marten and Barnes (*1979*) reviewed the literature concerning the development, modification, and application of in vitro rumen fermentation method for estimating forage quality. Our objectives will be to provide some general conclusions regarding the techniques, modifications, and source of errors.

Technique

One of the primary accomplishments of the NC-64 North Central Regional Research Committee in the United States as reported by Marten and Barnes (*1979*) was the recommendation of an in vitro rumen fermentation system that would estimate in vivo digestibility. We will summarize the methods outlined by Marten and Barnes as having potential for use in new laboratories initiating an in vitro system or for ongoing laboratories to compare with their current procedure. The methods were not claimed to be superior to numerous others published but to have yielded reproducible in vitro dry matter and organic matter results, as verified by collaborative trials. They have also produced results which will serve as adequate calibration for NIRS prediction.

Modification of the two-stage Tilley and Terry method. The Tilley and Terry (*1963*) method found great popularity among forage scientists, as attested to by its frequent use and citation. Marten and Barnes (*1979*) suggested modifications which include the use of a 250-mg rather than a 500-mg sample, the addition of urea to the bicarbonate buffer, or the use of a 24-hour instead of a 48-hour second stage (acid-pepsin) digestion. These modifications allow completion of in vitro runs within a normal work week.

In addition to the modifications listed above, Van Soest et al. (*1966*) proposed the use of a neutral detergent (ND) solution to solubilize part of the dry matter residue from the first stage. The acid-pepsin stage simulates in vivo breakdown of feed and microbial protein by digestive enzymes of the abomasum in the ruminant. The ND solution extracts cell solubles and leaves the undigested cell wall residue. A greater amount of dry matter is solubilized by the ND solution than by acid-pepsin. This has been attributed to the solubilization of the bacterial cell wall and other endogenous products by the ND solution. The undigested residue should represent the truly undigested matter of feed origin. Thus, the two-stage in vitro procedure with ND has been proposed for estimating the true digestibility of forages rather than the apparent digestibility (Van Soest et al. *1966*). The product of cell solubles × 0.98 is added to the ND fiber digested to obtain true digestibility.

Direct acidification method. The direct acidification method is a major modification of the Tilley and Terry (*1963*) method in that it employs direct acidification, without centrifugation, at the end of stage one. This is made possible by substitution of a phosphate buffer (with or without urea) for the original bicarbonate buffer. This substitution greatly saves time because it precludes the excessive frothing that occurs with bicarbonate buffer. It also includes the use of a 24-hour acid-pepsin stage.

Fungal enzyme method. Marten and Barnes (*1979*) presented a review of the literature on the fungal enzyme method, including the procedures and their correlations with in vivo and in vitro digestibilities of forages reported between 1973 and 1979. They reached the following conclusions regarding the use of fungal enzymes:

1. Fungal cellulases were often able to predict the digestibilities of forages nearly as well as in vitro rumen fermentation;
2. Fungal cellulases appeared to be more sensitive to forage species variations than did rumen inocula;
3. The use of ND pretreatment of substrates followed by a treatment with a standardized potent cellulase solution as described by Roughan and Holland (*1977*) in order to complete cellulase digestion needed confirmation;
4. Because some forage species responded differently than others to cellulase enzymes, standard samples of each species with known digestibilities needed to be included in each cellulase assay;
5. The activity of the selected cellulase needed to be measured before routine use; and
6. Because cellulases may vary greatly in their capacity to digest forage fiber, further research was needed to standardize the activity of marketed cellulase preparations.

Nylon bag method. The nylon bag method consists of placing samples in bags of an indigestible material such as nylon, Dacron, or silk within the rumen in situ.

The rate and extent of digestion are measured by the loss of dry matter (DM) from the substrate after a specific period of incubation. The method is subject to considerable variability and is difficult to standardize. Sources of variation include size and type of bags; cloth mesh size, sample size, and fineness of grind; number of samples per trial; diet of host animal; method, location, and time of suspension in the rumen; and method of rinsing the bags after removal from the rumen. Variability may be reduced by leaving the bags in the rumen for longer periods (Neathery *1969*); use of a large sample size (10 g), use of a large number of samples per trial (up to 48), and allowing the bags to move about freely within the rumen.

Sources of Variation

The errors associated with the in vitro methods may be described as either random or predicted. Random errors involve those factors contributing to the variability in the in vitro results. Predicting errors result from the failure of the in vitro data to estimate the in vivo parameters.

Random errors. The magnitude of the errors reflecting the precision of the in vitro methods is associated with the within and between trial variability. The failure to handle each sample in exactly the same way contributes to the random variation. Variation between trials is generally greater than variation within trials. The sources of variation that contribute to random errors are discussed below.

Fermentation vessel

Fermentation vessels that have been used include various glass and plastic containers as well as sealed, vented culture vessels. Sayre and Van Soest (*1972*) found that 122- × 28-mm glass centrifuge tubes provided lower in vitro dry matter digestibility (IVDMD) values that did 125-ml erlenmeyer flasks or 200- × 25-mm glass screw-capped tubes due to difficulty in providing proper agitation of the centrifuge tubes during fermentation. Minson and McLeod (*1972*) reported that vacuum infiltration of water into samples before inoculation to reduce substrate floating provided similar IVDMD values for polyethylene and polycarbonate centrifuge tubes. Thus, glass or plastic centrifuge tubes, vented to permit release of gas, are commonly used for most IVDMD methods. The use of the same type of tube within and between trials should prevent any bias.

Buffer-nutrient solution

A buffer-nutrient solution is used to control pH during fermentation and to supply nutrients for the rumen micro-organisms. The majority of in vitro methods for forage quality evaluations use artificial sheep saliva as described by McDougall (*1948*). The amount of buffer-nutrient solution does not appear to be as important as the amount of inoculum; however, a uniform ratio of substrate, inoculum, and buffer-nutrient solution is desirable to minimize trial to trial variation.

Inoculum source, processing, and amount used

The single greatest source of uncontrolled variation in any in vitro method is the inoculum. Standard forage samples of known digestibility should be included in each trial. These standards permit the adjustment of values for within and between trial variability and to determine when an entire trial should be discarded (Marten and Barnes *1979*). Viewing the inoculum with a light microscope for micro-organism activity is the simplest technique for determining if it is viable. However, a more reliable measure of rumen inoculum activity is needed.

The most accurage estimates of in vivo measurements are obtained from inoculum collected from an animal fed a forage similar to that being studied in vitro. This is particularly true when evaluating grasses and legumes (Bowden and Church *1962*). Researchers generally agree that donor animals should not receive grain in their diets to achieve best in vitro digestion with minimum variability. In addition, if the donor animal is fed hay only, the pH of the collected rumen inoculum should be within the recommended limits of 6.7 to 6.9.

Attempts have been made to improve the uniformity of inoculum by various processing methods. However, straining of collected rumen fluid through four layers of cheesecloth with no unnecessary delays, maintaining temperature at 38.5 or 39° C, and CO_2 gassing over inoculated substrates before stoppering of the vessel will yield satisfactory results. The amount of inoculum has little effect on in vitro results, provided its ratio with the substrate and nutrient buffer solution is held constant. If more inoculum is added without the appropriate increase in the buffer, the rate of digestion may increase and the differences between samples will be smaller (McLeod and Minson *1969*). Even though the amount of rumen inoculum has little effect on IVDMD values, the amount of indigestible material present in the inoculum can signficantly affect the within trial variability. Therefore, inoculum blanks containing buffer solution and rumen fluid are processed through both incubation stages. Marten and Barnes (*1979*) recommended six inoculum blanks interspersed throughout the forage samples. The DM residue is averaged from all blanks and used in calculating IVDMD values. We have found that 1 standard and 1 blank for each 20–22 tubes of samples is more appropriate. If one blank per rack of tubes is used, then each rack can be corrected within for the blank residue.

Sample preparation and size

Drying and grinding procedures influence IVDMD values much as they do other forage quality assays. Freeze-dried or fresh forages generally have higher IVDMD values than oven-dried samples. Oven-drying at excessively high temperatures produce indigestible artifacts. Usually, drying at 65° C or less is recommended. Freeze-drying is recommended when maintaining the morphology and chemical integrity of the samples is desired.

The particle size of ground forage samples is dependent upon the type of mill, the mill screen size used, the sharpness of the blades, and the moisture content of the sample at the time of grinding. Grinding disrupts the cell wall, enabling the microbial enzymes to penetrate into portions of the plant tissue from which they are normally excluded. Thus, the smaller the particle size, the greater the potential disappearance due to microbial degradation. Generally, a fine grind of 0.5-mm particle size is highly desirable, but 1-mm particle size is also satisfactory (Marten and Barnes *1979*). Numerous researchers have found that sample size can at times influence IVDMD values. However, variation can be controlled if the concentrations of buffer and rumen inoculum are maintained in constant ratio with the amount of substrate. Although the Tilley and Terry (*1963*) method calls for a 500-mg sample, many laboratories have successfully utilized 250-mg samples. One major problem with a smaller sample size is the variation resulting from systematic errors during weighing and transfer of sample. In addition, small sample size increases subsampling error.

Predicted errors. The second type of error inherent in the procedure is associated with the errors of predicting in vivo digestibility from in vitro estimates. Several potential problems can be theorized. The in vitro culture tube system is closed whereas the animal in the in vivo system is continually adding saliva, enzymes, and minerals and removing end products of digestion. Thus, depletion of nonfeed nutrients and cofactors and accumulation of digestion end products may limit digestion in the in vitro system. Continous culture systems have been developed to overcome some of these problems but are totally unsuitable for larger numbers of samples. Further errors involve metabolic products that are in the feces of animals used for in vivo trials and that are counted as undigested feed residue. These factors are not present in the in vitro system. The success of the in vitro system may revolve around the several errors which cancel each other out.

One error that has often been observed is that in vitro systems often underestimate in vivo digestibility of low quality forages and overestimate that of high quality forages. The reason is that the digestion period in the in vitro systems, unlike that in the in vivo systems, is fixed (48 hours). In the latter systems, undigested residues reside in the rumen for variable periods. Also, forages differ in rate of digestion. Typically, high quality forages have shorter residence times in the rumen, and the microbes in the in vitro systems have longer to solubilize the material; thus, overestimation results. The opposite is true for lower quality materials such as tropical grasses.

With all of their shortcomings, in vitro methods for determining digestibility of forages have helped considerably to increase our awareness of the importance of and improvement in forage quality.

References

Akin, D. E., D. Burdick, and H. E. Amos.
1974 Comparative degradation of Coastal bermudagrass, Coastcross-1 bermudagrass, and Pensacola bahiagrass by rumen microorganisms revealed by scanning electron microscopy. Crop Sci. 14:537-546.

Anderson, D. M., D. A. Christensen, and G. I. Christison.
1977 Voluntary intake and digestibility of slough hay and dehydrated alfalfa pellets by rams and steers. Can. J. Anim. Sci. 57:609-611.

AOAC.
1975 Official methods of analysis. 12th ed. Association of Official Agricultural Chemists, Washington, DC.

Balch, C. C.
1971 Proposal to use time spent chewing as an index of the extent to which diets for ruminants possess the physical property of fibrousness characteristic of roughages. Brit. J. Nutr. 26:383-392.

Barnes, R. F.
1965 Use of in vitro rumen fermentation technique for estimating forage digestibility and intake. Agron. J. 57:213-216.

Barnes, R. F.
1973 Laboratory methods of evaluating feeding value of herbage. *In* G. W. Butler, and R. W. Bailey, eds., Chemistry and biochemistry of herbage. p. 179-214. Academic Press Inc., Ltd., London.

Beardsley, D. W.
1964 Symposium on forage utilization. Nutritive value of forage as affected by physical form. Part II. Beef cattle and sheep studies. J. Anim. Sci. 23:239-245.

Bines, J. A., S. Suzuki, and C. C. Balch.
1969 The quantitative significance of long-term regulation of food intake in the cow. Brit. J. Nutr. 23:695-704.

Blaxter, K. L., N. McC. Graham, and F. W. Wainman.
1956 *Some* observations of the digestibility of food by sheep, and on related problems. Brit. J. Nutr. 10:69-91.

Blaxter, K. L., F. W. Wainman, and R. S. Wilson.
1961 The regulation of food intake by sheep. Anim. Prod. 3:51-61.

Bowden, D. M., and D. J. Church.
1962 Artificial rumen investigations. I. Variability of dry matter and cellulose digestibility and production of volatile fatty acids. J. Dairy Sci. 45:972-979.

Brody, S.
1945 Bioenergetics and growth. Hafner Pub. Co., Inc., New York.

Campling, R. C., M. Freer, and C. C. Balch.
1961 Factors affecting the voluntary intake of food by cows. 2. The relationship between the voluntary intake of roughages, the amount of digesta in the reticulo-rumen and the rate of disappearance of digesta from the alimentary tract. Brit. J. Nutr. 15:531-540.

Coleman, S. W., O. Neri-Flores, R. J. Allen, and J. E. Moore.
1978 Effect of pelleting and of forage maturity on quality of two sub-tropical forage grasses. J. Anim. Sci. 46:1103-1112.

Colovos, N. F., J. B. Hotter, R. M. Koes, and others.
1970 Digestibility, nutritive value and intake of ensiled corn plant (*Zea mays*) in cattle and sheep. J. Anim. Sci. 30:819-824.

Conrad, H. R.
1966 Symposium on factors influencing the voluntary intake of herbage by ruminants: physiological and physical factors limiting feed intake. J. Anim. Sci. 25:227-235.

Conrad, H. R., A. D. Pratt, and J. W. Hibbs.
1964 Regulation of feed intake in dairy cows. I. Change in importance of physical and physiological factors with increased digestibility. J. Dairy Sci. 27:54-62.

Crampton, E. W.
1957 Interrelations between digestible nutrient and energy content, voluntary dry matter intake, and the overall feeding value of forages. J. Anim. Sci. 16:546-552.

Crampton, E. W, E. Donefer, and L. E. Lloyd.
1960 A nutritive value index for forages. J. Anim. Sci. 19:538-544.

Evans, E. M., J. F. Potter.
1984 The reproducibility of in vivo estimates of digestibility and voluntary digestible organic matter intake of grass varieties by sheep. Grass and Forage Sci. 39:101-106.

Foot, J. Z.
1972 A note on the effect of body condition on the voluntary intake of dried grass wafers by Scottish Blackface ewes. Anim. Prod. 14:131-134.

Forbes, J. M.
1977 Interrelationsjips between physical and metabolic control of voluntary intake in fattening pregnant and lactating mature sheep: a model. Anim. Prod. 24:91-101.

Forbes, E. B., R. F. Elliot, R. W. Swift, and others.
1946 Variations in determinations of digestive capacity of sheep. J. Anim. Sci. 5:298-305.

Forbes, J. M., A. A. El Shahat, R. Jones, and others.
1979 The effect of daylength on the growth of lambs. 1. Comparisons of sex, level of feeding, shearing and breed of sire. Anim. Prod. 29:33-42.

Goering, H. K., and P. J. Van Soest.
1970 Forage fiber analyses (apparatus, reagents procedures and some applications). U.S. Dept. Agric., Agric. Handbook No. 379.

Heaney, D. P.
1969 Voluntary intake as a component of an index to forage quality. *In* Proc. Nat. Conf. Forage Qual. Eval. and Util. Prog. p. C1-10. Lincoln, NE.

Heaney, D. P.
1979 Sheep as pilot animals. *In* W. J. Pigden, C. C. Balch and M. Graham, eds., Standardization of analytical methodology for feeds. p. 44-48. International Development Research Center, Ottawa, Canada.

Heaney, D. P., G. E. Pritchard, and W. J. Pigden.
1968 Variability in *ad libitum* forage intakes by sheep. J. Anim. Sci. 27:159-164.

Homb, L. E.
1963 Qualitative estimation of forage plant with particular references to methods of study of the digestibility in vitro and in vivo. *In* Proc. Europ. Grassl. Symp., p. 29-36. British Grassland Society, Hurley, United Kingdom.

Hungate, R. E.
1966 The rumen and its microbes. Academic Press, New York.

Johnson, R. R.
1963 Symposium on microbial digestion in ruminants: in vitro rumen fermentation techniques. Anim. Sci. 22:792-800.

Karue, C. N., J. L. Evans, and A. D. Tillman.
1973 Voluntary intake of dry matter by African Zebu cattle. Quality of feed and the reference base. J. Anim. Sci. 36:1181-1185.

Kilmer, L. H., P. J. Wangsness, E. M. Kesler, and others.
1979 Voluntary intake and digestibility of legume and grass diets fed to lactating cows and growing wethers. J. Dairy Sci. 62:1272-1277.

Kleiber, M.
1961 The fire of life—an introduction to animal energetics. John Wiley and Sons, Inc., New York.

Laredo, M. A., and D. J. Minson.
1973 The voluntary intake, digestibility and retention time by sheep of leaf and stem fractions of five grasses. Aust. J. Agric. Res. 24:875-888.

Lippke, H.
1980 Forage characteristics related to intake, digestibility and gain by ruminants. J. Anim. Sci. 50:952-961.

Lucas, H. L., Jr., W. W. G. Smart, Jr., M. A. Cipolloni,
1961 and H. D. Gross. Relations between digestibility and composition of feeds and foods. S-45 Report. North Carolina State College, Raleigh.

Marten, G. C., and R. F. Barnes.
1979 Prediction of energy digestibility of forages with in vitro fermentation and fungal enzyme systems. *In* W. J. Pigden, C. C. Balch, and M. Graham, eds., Standardization of analytical methodology for feeds, p. 49-60. International Development Research Center, Ottawa, Canada.

McDougall, E. I.
1948 Studies on ruminant saliva. I. The composition and output of sheep saliva. Biochem. J. 43:99-109.

McLeod, M. N., and D. J. Minson.
1969 Sources of variation in the in vitro digestibility of tropical grasses. J. Br. Grassl. Soc. 24:244-249.

McLeod, M. N., and D. J. Minson.
1974 Predicting organic-matter digestibility from in vivo and in vitro determinations of dry matter digestibility. J. R. Grassl. Soc. 29:17-21.

McLeod, M. N., and D. J. Minson.
1978 The accuracy of the pepsin-cellulase technique for estimating the dry matter digestibility *in vivo* of grasses and legumes. Anim. Feed Sci. Tech. 3:227-287.

Milford, R., and D. J. Minson.
1965 Intake of tropical pasture species. *In* Proc. 9th Int. Grassl. Congr., p. 814-822. Sao Paulo, Brazil.

Minson, D. J., K. L. Butler, N. Grummitt, and D.P. Law.
1983 Bias when predicting crude protein, dry matter digestibility and voluntary intake of tropical grasses by near-infrared reflectance. Anim. Feed Sci. Tech. 9:221-237.

Minson, D. J., and J. L. Cowper.
1977 Equipment for automatically feeding forage to cattle at hourly intervals and removing uneaten feed. J. Anim. Sci. 44:814-817.

Minson, D. J., C. E. Harris, W. F. Raymond, and R. Milford.
1964 The digestibility and voluntary intake of S 22 and H.1 ryegrass, S 170 tall fescue, S 48 timothy, S 125 meadow fescue and germinal cocksfoot. J. Brit. Grassl. Soc. 19:298-305.

Minson, D. J., and M. M. McLeod.
1972 The in vitro technique: its modification for estimating digestibility of large numbers of tropical pasture samples. CSIRO, Australia, Dir. Trop. Past. Tech. Paper No. 8, p. 1-15.

Minson, D. J., and R. Milford.
1967 The voluntary intake and digestibility of diets containing different proportions of legumes and mature pangolagrass (*Digitaria decumbens*). Aust. J. Agric. Anim. Husb. 7:546-551.

Monson, W. G., R. S. Lowrey, and I. Forbes, Jr.
1969 *In vivo* bag vs. two-stage in vitro digestion: comparison of two techniques for estimating dry-matter digestibility of forages. Agron. J. 61:587-589.

Montgomery, M. J., and B. R. Baumgardt.
1965 Regulation of food intake in ruminants. 1. Pelleted rations varying in energy concentration. J. Dairy Sci. 48:569-574.

Moore, J. E.
1977 Southern forages in the new hay standards. *In* Proc. 34th South. Past. and Forage Crop Impr. Conf. p. 129.

Moore, L. A.
1964 Symposium on forage utilization. Nutritive value of forage as affected by physical form. I. Several principals involved with ruminants and effect of feeding pelleted or wafered forage to dairy cattle. J. Anim. Sci. 23:230-238.

Neathery, M. W.
1969 Dry matter disappearance of roughage in nylon bags suspended in the rumen. J. Dairy Sci. 52:74-78.

Norris, K. H., R. F. Barnes, J. E. Moore, and J. S. Shenk.
1976 Predicting forage quality by infrared reflectance spectroscopy. J. Anim. Sci. 43:889-897.

Pigden, W. J.
1969 Laboratory analyses of herbage used to predict nutritive values. *In* J. B. Campbell, ed., Experimental methods for evaluation herbage, p. 52-72. Ottawa, Public 315, Can. Dep. Agric., Queens Center.

Purser, D. B., and R. J. Moir.
1966 Rumen volume as a factor involved in individual sheep differences. J. Anim. Sci. 25:509-515.

Putnam, P. A., and R. E. Davis.
1961 Effect of feeding pelleting complete rations on lactating cows. J. Dairy Sci. 44:1465.

Raymond, W. F.
1969 The nutritive value of forage crops. Adv. Agron. 21:1-108.

Raymond, W. F., C. E. Harris, and V. G. Harker.
1953 Studies on the digestibility of herbage. I. Technique of measurement of digestibility and some observations of factors affecting the accuracy of digestibility data. J. Br. Grassl. Soc. 8:301-314.

Reid, R. L.
1966 Forage evaluation. *In* Proc. 12th Annual Farm Seed Conference, p. 7-17.

Roughan, P. G., and R. Holand.
1977 Predicting *in vivo* digestibility of herbages by exhaustive enzyme hydrolysis of cell walls. J. Sci. Food Agric. 38:1057-1064.

Sayre, K. D., and P. J. Van Soest.
1972 Comparison of types of fermentation vessels for an *in vitro* artificial rumen procedure. J. Dairy Sci. 55:1496-1498.

Schneider, B. H., and W. P. Flatt.
1975 The evaluation of feeds through digestibility experiments. p. 6-51. The University of Georgia Press, Athens.

Shelton, D. C., and R. L. Reid.
1960 Measuring the nutritive value of forages using *in vitro* rumen techniques. Int. Grassl. Congr. Proc. 8:524-528.

Shenk, J. S., I. Landa, M. R. Hoover, and M. O. Westerhaus.
1981 Description and evaluation of a near infrared reflectance spectrocomputer for forage and grain analysis. Crop Sci. 21:355-358.

Sudweeks, E. M., M. E. McCullough, L. R. Sisk, and S. E.
1975 Law. Effects of concentrate type and level and forage type on chewing time of steers. J. Anim. Sci. 41:219-224.

Taylor, St. C. S., and G. B. Young.
1968 Equilibrium weight in relation to food intake and genotype in twin cattle. Anim. Prod. 10:393.

Tilley, J. M. A., and R. A. Terry.
1963 A two-stage technique for the *in vitro* digestion of forage crops. J. Brit. Grassl. Soc. 18:104-111.

Torrie, R. A., de la.
1974 Micro-histological characteristics of three warm-season grasses in relation to forage quality. Ph.D. dissertation. [On file at the University of Florida, Gainsville.]

Troelsen, J. E., and F. W. Bigsby.
1964 Artificial mastication—a new approach for predicting voluntary forage consumption by ruminants. J. Anim. Sci. 23:1139-1142.

Van Soest, P. J.
1965 Symposium on factors influencing the voluntary intake of herbage by ruminants: voluntary intake in relation to chemical composition and digestibility. J. Anim. Sci. 24:834-843.

Van Soest, P. J., D. R. Mertens, and B. Denium.
1978 Preharvest factors influencing quality of conserved forage. J. Anim. Sci. 47:712-720.

Van Soest, P. J., R. J. Wine, and Moore, L. A.
1966 Estimation of the true digestibility of forage by the in vitro digestion of cell walls. *In* Proc. 10th Int. Grassl. Congr., p. 438-441. Helsinki, Finland.

Ventura, M., J. E. Moore, O. C. Ruelke, and D. E. Franke.
1975 Effect of maturity and protein supplementation on voluntary intake and nutrient digestibility of Pangola digitgrass hays. J. Anim. Sci. 40:769-774.

Waldo, D. R.
1969 Factors influencing the voluntary intake of forages. *In* Proc. Nat. Conf. Forage Qual. Eval. and Util. Prog., p. E1-21. Lincoln, NE.

Ward, R. G., G. S. Smith, J. D. Wallace, and others.
1982 Estimates of intake and quality of grazed range forage by near infrared-reflectance spectroscopy. J. Anim. Sci. 54:399-402.

Waldo, D. R., L. W. Smith, and E. L. Cox.
1972 Model of cellulose disappearance from the rumen. J. Dairy Sci. 55:125-129.

Welch, J. G., and A. M. Smith.
1969 Influence of forage quality on rumination time in sheep. J. Anim. Sci. 28:813-818.

1. Protocol for NIRS Calibration: Sample Selection and Equation Development and Validation

W. R. Windham, D. R. Mertens, and F. E. Barton II

Introduction

Calibrating a near infrared reflectance spectroscopy (NIRS) instrument involves using a set of samples to develop a mathematical relationship between spectra generated by the instrument and values obtained from a laboratory reference method. Calibrations range from empirical (where the mathematical relationship is unknown and must be estimated from a set of samples) to analytical (where the form of the mathematical relationship is known ahead of time). The NIR spectra and chemistry of agricultural products are complex; little is known about the mathematical relationships between them. Use of NIRS for forage analysis has relied on the empirical approach. NIRS instruments can be successfully calibrated for forage quality constituents by the following steps:

1. Selection of a calibration sample set.
2. Collection of NIR spectral data.
3. Analysis of the calibration sample set by a primary reference method.
4. Calibration of the instrument by generating regression equations using various mathematical treatments of spectral data and selection of the best equation that relates reference method analysis to NIR spectral data.
5. Validation of the final equation with a similar sample set (closed population) or a randomly acquired new sample set (open population).

The manner in which steps 1-4 are conducted has long been recognized as the basis for successful analysis of forage quality by NIRS (Shenk et al. *1979*). The literature abounds with information regarding NIRS analysis of forage quality; however, specific protocols for successful equation generation and selection are variable. Generally, researchers state that "calibration equations were chosen by a combination of statistics from calibration development (small standard error of calibration, large R^2 and F statistic of the last entered wavelength)," and that equations were "validated by the actual concentration compared to the predicted concentration using r^2, bias, and standard error of analysis" (Windham et al. *1987*). Although these statements are descriptive, they are incomplete considering the large number of initial equations from which the analyst must select the "best" equation for use.

In addition, there are considerable differences in the literature regarding the statistical expression of equation validation. The standard error of NIRS determinations has been expressed as standard error (SE) of prediction (SEP), performance (SEP), validation (SEV), and analysis (SEA). All these statistical expressions have had the same meaning and were calculated as described by Abrams et al. (*1987*). The National NIRS Forage Research Project Network now endorses the following terminology to clarify standard error expressions. The standard error of calibration is expressed as SEC and defined as $[\Sigma(X_i - Y_i)^2/(N - p - 1)]^{0.5}$, where X_i is the value determined by conventional analytical methods, Y_i is the value determined by NIRS, N is the number of samples, and p is the number of terms in the calibration equation. This is equivalent to the standard error of multiple regression. The standard error of equation performance when selecting the final equation during the calibration procedure is expressed as standard error of selection (SES). The standard error of analysis of samples used in validating or monitoring the use of NIRS is standard error of performance (SEP). Both SES and SEP are defined as $[\Sigma(X_i - Y_i)^2/(N - 1)]^{0.5}$, where X_i, Y_i and N are as previously defined (except that X_i and Y_i are from different populations). Both SES and SEP are separated into two components, bias and unexplained error. Bias corrected standard errors are expressed as SES(C) and SEP(C) depending on where the statistic is used. In the monitoring system, SEP(C) is expressed as unexplained error. Standard error of the laboratory (SEL) is the standard error of variance between replicates analyzed by the reference method. The SEL is defined as $[\Sigma_i[\Sigma_j(X_{ij} - \bar{X}_i)^2/(R - 1)]/N]^{0.5}$, where X_{ij} *is the* jth replicate on the ith sample, $\bar{X}_i$ is the reference method mean value of all replicates of an ith sample, R is the number of replicates, and N is the number of samples.

The objectives of this supplement are to provide general and specific protocols for selection of calibration and validation samples and for equation generation, selection, and validation.

Calibration Sample Selection

Calibration is accomplished with a subset of samples from a larger population. All factors affecting the NIR spectra must be represented in the calibration set. These factors include sample physical and chemical characteristics, method of sample preservation and processing, and instrument and sample environment (that is, temperature and humidity). These factors can be reduced by controlling sample processing as well as instrument and sample environment. Most accurate NIRS calibration will occur when the major factor affecting the spectra is the variability in sample physical and chemical characteristics. (See "Equation Selection," p. 26, for further examples.) Sampling theory indicates that randomly selecting a predetermined number of samples from the initial sample set may be the best way to obtain a calibration sample set that will yield unbiased NIRS equations. Structured sample selection based on known differences in sample type,

preservation, and growth environment may be appropriate in a closed population, especially if these factors are related to differences in the physical and chemical characteristics of the samples.

Alternatively, samples may be selected based on spectral characteristics. A software program (SUBSET) has been developed by the National NIRS Forage Research Project Network as an aid to calibration sample selection. The program SUBSET has two functions. First, it can be used to eliminate samples with similar spectra from a file to be used for calibration. Second, it can be used to determine whether new samples are sufficiently different from samples already in the calibration file to warrant being added. The underlying assumption is that a calibration should include as few samples as possible. If the population is closed or narrow, the program will eliminate redundant samples. If the population is open or broad based, the program will eliminate redundant samples from the starting population and screen potential samples for calibration update.

The technique used by SUBSET is the "neighborhood" concept. The concept involves correlation among spectra in the file. Two samples are in the same neighborhood if the correlation between their spectra is greater than a predefined value. Correlations are calculated between the spectra of all possible pairs of samples. Correlation was chosen as the statistic for comparison because it is independent of multiplicative factors such as particle size differences, can be calculated rapidly, and does not require reference analysis on the samples. The segment limits for comparison are usually set from 1108 to 1848 nm and from 1968 to 2478 nm for agricultural products. This setting eliminates most of the effect of water on the spectra so that samples will not be classified into groups based on differences in water concentration.

The tolerances of correlation coefficients for spectral selection of calibration samples are set by default in the program (0.950, 0.997 r^2) or set by the operator based on experience. The first value, 0.950, is used in the connectedness test. Any sample whose largest correlation with any other sample in the sample population is less than 0.950 is considered not connected (that is, dissimilar) to the rest of the samples, and it will not be selected for the calibration set. Any sample with a correlation greater than 0.997 with an already selected sample is considered too similar to be included in the calibration set. Any sample whose highest correlation with an already selected sample is less than 0.997 but greater than 0.950 is considered to be a sample with variation that is needed in the calibration set.

Although these concepts are easy to understand, applying them properly to a population of samples is sometimes difficult. If a population of samples to be used for calibration is homogeneous, the 0.997 limit for similar samples may not result in enough samples to develop a reasonable calibration. In this case, the 0.997 limit should be increased in order to select additional samples for calibration. Likewise, as the population is broadened, the value may need to be reduced to keep the number of samples used for calibration to a minimum.

The number of samples is important in calibration. Small calibration sets may not accurately represent unknown populations. The recommended minimum number of calibration samples for any forage quality parameter in closed populations is about 50. The maximum number of samples is dependent on the number of within-population factors that affect the NIR spectra, but rarely do benefits occur from exceeding 250. Table 1 contains the number of samples used by various researchers for NIRS calibration of four forage-quality parameters. On the average, 75

Table 1.—*Number of samples* (N) *and standard error of calibration (SEC) associated with NIRS calibration of 4 forage-quality parameters*

	Quality parameters							
	CP[a]		NDF[b]		ADF[c]		IVDMD[d]	
Reference	*N*	SEC	*N*	SEC	*N*	SEC	*N*	SEC
Norris et al. *1976*	87	0.74	87	2.74	87	1.64	87	2.64
Shenk and Barnes *1977*	117	.78	---	---	---	---	117	1.96
Shenk et al. *1978*	---	---	117	2.16	117	1.96	---	---
Marten et al. *1983*	48	.51	48	1.40	48	.61	48	2.30
Parnell and White *1983*	108	.70	---	---	---	---	49	1.20
Martin et al. *1985*	123	.57	65	1.83	65	1.56	---	---

[a] Crude protein. [b] Neutral detergent fiber. [c] Acid detergent fiber. [d] In vitro dry matter digestibility.

samples were used for calibration equations. In general, they were narrow based (that is, closed population) equations developed for finite populations, and they had limited value beyond their use for those populations. (See "Equation Selection," p. 26, for further discussion of narrow- vs. broad-based populations.) Conversely, Abrams et al. (*1987*) reported the development of broad-based forage-quality calibration equations from 567 hay samples. Calibration subsets of 70, 100, 151, and 200 samples were randomly selected from 467 samples that represented 5 species, 5 years, and 3 cuttings. Samples were preserved by oven drying (50, 65, or 80° C), or freeze drying, or 1, 2, or 3 days of field drying. Calibration subsets were selected either randomly or by utilizing statistical clustering, in which samples of like spectra were placed in clusters. The remaining 100 samples were used as a validation population. Equations for CP, NDF, ADF, and IVDMD within calibration subsets with the smallest SEC were applied to the validation population. The standard error of performance (SEP) and bias were reduced curvilinearly ($P < 0.05$) for all assays with increasing size of the calibration set. Calibration subsets of 70 samples resulted in the largest SEP for all forage-quality parameters. Calibration subsets of 151 samples (32 percent of those available) were necessary to obtain an SEP and bias as low as those acquired with all 467 samples. No differences in SEP and bias were due to the selection method.

From these results several general guidelines can be derived regarding selection of calibration samples:

1. Calibration samples should be randomly selected from the population to be analyzed, provided all major sample types are represented by the random selection. Alternatively, samples can be selected on the basis of spectral features in which the minimum number of samples are chosen to represent all the spectral variability in the total population (that is, subset).

2. With narrow-based populations, less than 100 samples (minimum 50) are usually adequate for calibration.

3. With broad-based populations, 150 samples or more are necessary for calibration.

4. In both types of populations, emphasis should be on selection of samples that are representative of major variables affecting sample physical and chemical characteristics (that is, growing environments, stage of maturity, plant species, method of preservation, particle size and grinding, and water concentration).

Reference Methods

No attempt is made here to discuss the various reference methods used for NIRS calibration. Some information on these methods is in "Animal Response Prediction" (p. 32). The most important aspect of the reference data used in the calibration process is to maintain a high level of precision. The maximum allowable SEL within a laboratory are $\leqslant$ 0.3, 0.6, 1.2, 1.5, and 2.0 for DM (dry matter), CP, ADF, NDF, and IVDMD, respectively. However, a more acceptable laboratory error is one-half the maximum allowable SEL. Reference methods may not be well defined chemically, but if they are followed closely, they can be precise and repeatable.

Protocol

Sample Preparation

Dry samples in a vacuum, forced-air, or microwave oven; or freeze dry them or leave them in the field to dry. Any method involving controlled temperature should use the same temperature and time of drying for all samples, both calibration and unknown.

Grind samples for NIRS calibration and analysis in a cyclone mill or a Wiley mill through a 1-mm screen. Grind all subsequent samples to be analyzed the same way as those in the calibration set. The mills should be cleaned between samples to minimize cross-contamination. (See "Sample Preparation," p. 23, for further details.) For closed populations, samples should be processed over a short period and under the same conditions of preservation and grinding. Open population samples should be processed over several weeks to account for systematic changes in sample processing that can occur.

Mix milled samples well, and place four random portions in four quadrants of the NIRS cell to ensure that portions of different subsamples are scanned. Continue to take random test portions until the NIRS sample holder is level full and scrape off any excess. Take the rubber or cardboard back and press it into the holder until it is tight and level. As a check, invert the NIRS sample holder and make certain the sample is firmly pressed against the window. If any abnormality is apparent, remove the back and repeat this procedure. Consistency in sample handling and preparation is crucial to the successful use of NIRS analysis.

Spectral Data Collection

To develop equations, collect reflectance (R) data (log 1/R) of the calibration population with the program SCAN. (See "Public Software," p. 18.) Develop robust equations for open populations by collecting reflectance data over days or weeks to ensure that differences in spectroscopic data due to changes in laboratory humidity and temperature are included in the calibration set. Select calibration samples randomly, and scan them in a random sequence.

Merge all files created into one calibration file with the program FILE.

Enter primary reference data from analysis of the calibration samples with the reflectance data of the calibration samples using the program DATA.

Preliminary Calibration Set Evaluation

Conduct a preliminary NIRS calibration with the program BEST to evaluate the reference method data using only one mathematical treatment (1,10,10,2 in nanometers or 1,5,5,1 in data points) and less than $N/10 + 1$ terms or wavelengths, where N = number of samples in the calibration file. BEST provides the option of splitting the calibration file into two subsets, one file for equation generation and one for equation selection. However, it should only be used if the samples are scanned in a random order. If the data are entered in an order associated with treatments, selecting every *i*th sample can result in a nonrandom, stratified sample set. The guidelines for splitting calibration files are given in table 2. The splits in table 2 result in 20 to 30 percent of the samples being used for equation selection during calibration. Usually 20 samples should be considered the minimum for equation selection. If too few are included in the selection subset, the statistics for selecting equations may be misleading. With 50 total calibration samples, an *i*th split of 3, and beginning at file position 1, 2, or 3, only 17, 18, or 16 samples, respectively, are used for selection. If less than 20 samples are used for equation selection in the preliminary calibration set evaluation, 2 or 3 calibrations should be done with a different selection subset each time to ensure that the "best" equations are generated for the selection step.

Table 2.—*Relationship between number of total calibration and equation selection samples during equation development*

Number of calibration samples	*i*th split[a]	Number of equation selection samples[b]
<50	None	None
50	3	16
100	3	33
100	4	25
150	3	50
150	4	37
200	4	50
200	5	40

[a] Option in program BEST for splitting the file into a subset for calibration and a subset for equation selection, where i = 1 to 5.

[b] Number of equation selection samples for each *i*th split will vary slightly depending on the file position of the sample from which the *i*th split occurs.

Obtain a printed copy of the preliminary calibration data. Compare the mean and standard deviation of the reference method data in the portion of samples used for equation generation with the mean and standard deviation of the reference method data in the equation selection sample set. If they differ by more than 20 percent, the random file split is not acceptable and the preliminary calibration should be rerun with a file split that results in a smaller difference. This ensures that the samples used for equation selection are comparable to those used for equation generation and assures that the file was appropriately split. Within the preliminary calibration set, equations with 1 through $N/10 + 1$ terms (wavelengths) will be obtained using the generation set and will be evaluated using the selection sample set.

For each constituent, find the equation where the standard error of selection corrected for bias (SES(C)) is minimum or a plateau value. Compare the SES(C) and the SEC for that equation. As a guideline, the values should be within 20 percent of each other. Examine the differences (residuals) between NIRS determinations (that is, generation and selection sets) and the reference method determination for samples with large t values. A large positive or negative t value indicates that the residual is greater than 2.5 times the standard error of a difference (SED) between the NIRS and reference determinations. Less than 5 percent of the total samples should be t outliers. If the same samples are outliers in every equation, the subsample analyzed by the reference method is not representative of that which was scanned, or the data from the reference method are inaccurate, or the outlier samples do not belong in the calibration population. Reanalyze the outlier samples by both methods. In small calibration populations, t outliers can make a large difference in the calibration statistics, but in a large population (over 100 samples), the effect may be negligible. The statistics can be improved by removal or reanalysis of outliers.

Next, evaluate the samples that are spectral outliers (that is, the samples' "H" statistic is large relative to the expected H statistic). A large (three and larger) H statistic indicates that the NIR spectrum for a given sample differs substantially from the NIR spectrum of the other calibration samples. Rescan such a sample and compare the rescan with the original scan. If the sample belongs in the population (that is, two scans agree), leave it in the calibration because it is an important sample. Satisfactory

results by this procedure indicate that the reference method data are accurate and that a good calibration can be obtained with the selected population.

Equation Generation

As preferred, use the same file split from the preliminary evaluation or another based on the guideline in table 2. Run the program BEST with a maximum number of terms or wavelengths of $N/10 + 1$ and the following mathematical treatments of log 1/R data:

Nanometers (nm)	*Data points (DP)*
1,10,10,2	1,5,5,1
1,20,10,2	1,10,5,1
1,30,10,2	1,15,5,1
2,20,20,2	2,10,10,1
2,30,20,2	2,15,10,1
2,40,20,2	2,20,10,1

Selection of Equations Within Mathematical Treatments

For each constituent, evaluate the standard error of calibration (SEC) and R^2 for each equation. The SEC will decrease progressively and the R^2 will increase as wavelengths or terms are added. Observe the SEC to assure that the better equations have values in the expected range (that is, DM < 0.6, CP < 0.9, ADF < 1.5, NDF < 2.0, and IVDMD < 3.0). The R^2 should generally be 0.9 or better. Recognize that a constituent with a small range of values may show lower correlation coefficients. (See "Animal Response Prediction," p. 32, for the relationships of standard deviation of data sets and R^2.)

Next, evaluate the SES(C) of each equation within each mathematical treatment using the equation selection sample set. The SES(C) is an indication of the performance of the equations on the selection sample set. For each mathematical treatment, find the equation where SES(C) is minimum or a plateau value. Unlike SEC, which must decrease with each additional term, SES(C) decreases only until overfitting of equations in the generation sample set becomes important and causes SES(C) to increase. Find the SEC in the generation sample set for that equation where SES(C) in the selection sample set has plateaued or reached a minimum. Again, as a guide, the values should be within 20 percent of each other. Select the "best" equation(s) within a mathematical treatment with the following set of guidelines:

1. The equation should have the lowest SES(C) and fewest terms or wavelengths.

2. Wavelengths in the equation should be greater than 40 nm apart.

3. No wavelengths should have an *F* statistic less than 10 for the regression coefficients. Start with a 1-term equation and stop evaluating equations after an equation contains coefficients with *F* value below 10.

4. The regression coefficient size should be inspected for the selected equations. As a guide, the coefficient for a derivative mathematical treatment should not be larger than 10,000. Coefficient size is a function of mathematical treatment. Narrow gaps, that is, 1, 20, 10, 2, will usually have larger coefficients because differences are a function of the smoothing band width.

5. The slope of the regression line relating the NIRS determinations to the primary reference values should be evaluated. Slopes should not exceed a range of 0.95 to 1.05. If slopes are markedly different from 1.0, high and low values will be especially biased.

Based on these guidelines, select one equation within each mathematical treatment.

Selection of Final "Best" Equation Among Mathematical Treatments

After the "best" equation from each mathematical treatment has been selected, summarize the results under the following headings:

Mathematical treatment
Number of wavelengths
SES(C)
SEC
F *test*
Regression coefficient

To select the "best" single equation for use in NIRS analysis, choose the equation with the lowest SES(C). If two or more equations have similar SESs (that is, difference is less than 10 percent), reevaluate the equation generation statistics for those equations. Choose the equation that has a combination of the fewest terms, an SEC similar to the SES(C), small regression coefficients, large r^2, and high *F* statistics associated with wavelengths identified for the constituents. When the "best" equation for each constituent has been selected, use the program BEST to fit that mathematical treatment and wavelengths on all calibration samples (that is, those combined for generation and selection) to derive the final equation that will be used in NIRS analysis.

Equation Validation

The final step in calibration is validation of the selected equation with samples not included in the original calibration (generation and selection) population. This step

is necessary to obtain an independent measure of equation accuracy expressed as standard error of performance (SEP).

From the unknown samples that have been analyzed using NIRS, randomly select 10 samples or 2 percent, whichever is larger. Analyze these samples using the reference method. Enter the data from the primary reference method for each validation sample using the program DATA.

Create a separate prediction file for these samples using the program PRE.

Compare the actual primary reference data with the predicted NIRS data using the program STAT. Typically the SEL of a reference method will be two-thirds the standard deviation of blind replicates. If the SEP for validation is within 2 times the SEL for the replicated primary reference method analysis, the final NIRS equation can be accepted for use, and the SEP for validation can be used as a reliable indication of the accuracy of the final NIRS equation. An acceptable relationship between NIRS analysis and the reference method is shown in figure 1. The slope is essentially 1, the intercept is 0, and the error, as indicated by scatter about the line, is reasonably low. If the SEP for validation is too large, then the source of the variation must be identified and removed, if possible, to obtain a satisfactory NIRS equation.

First, determine the relationship between SEP, SEP(C), and SEC. The SEP(C) statistic indicates the performance of the equation corrected for the differences (bias) between the final NIRS equation and the primary reference data. Figure 2 demonstrates the effect of a constant bias. Bias is defined as $\Sigma(X_i - Y_i)/N$. In this case, there is some factor that gives uniform biased values for NIRS analysis, resulting in a shift of the line relating NIRS to the reference method. This bias could be due to a change in the sample preparation and processing, primary reference method, or instrument since the calibration. However, if the SEP(C) is substantially lower than SEP for the validation samples and within 2 times the duplicate SEL for the reference method, the NIRS equation can be used by correcting for bias.

Unacceptable relationships between NIRS and the reference method analysis are shown in figures 3 and 4. Both graphs exhibit a large unexplained error (SEP(C)). Figure 3 demonstrates the effect of changes in slope. Samples with low reference method values are overestimated by NIRS, and samples with high reference method values are underestimated. Figure 4 demonstrates an increase in SEP(C), although the intercept remains 0 and the slope 1. A loss of accuracy has taken place, as indicated by the high random scatter about the line. A high SEP(C) may occur for three reasons: (1) The current wavelengths in the equation and resulting regression coefficients cannot be used to reliably estimate the

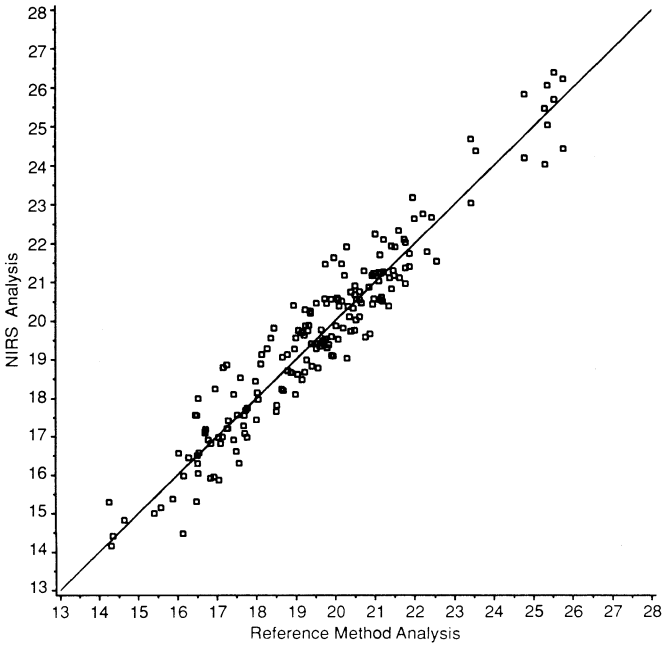


Figure 1.—An acceptable scatter relationship between NIRS analysis and the reference method.

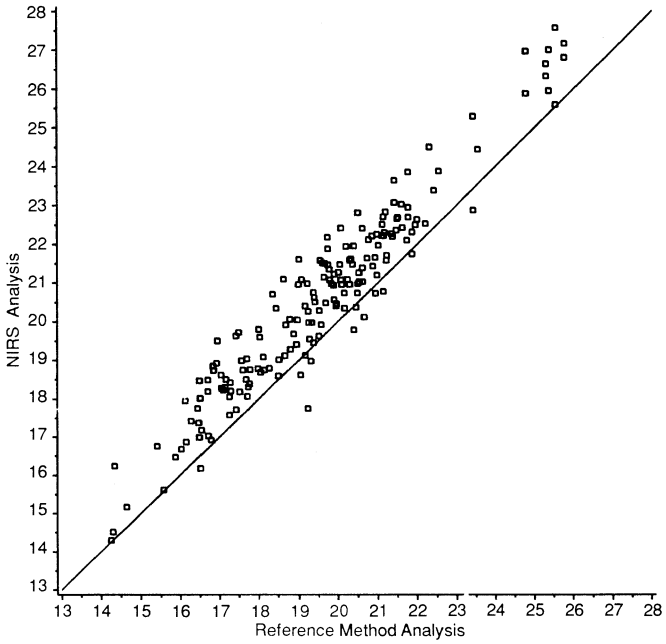


Figure 2.—An unacceptable bias effect between NIRS analysis and the reference method.

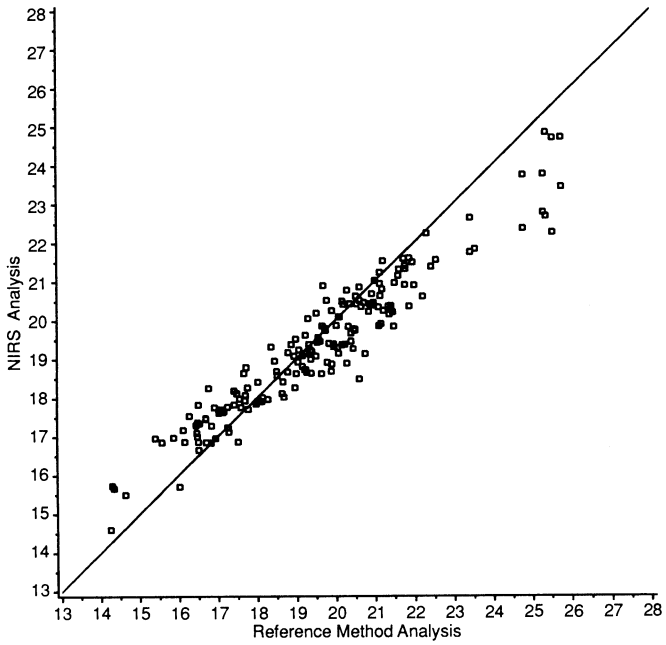


Figure 3.—An unacceptable slope effect between NIRS analysis and the reference method.

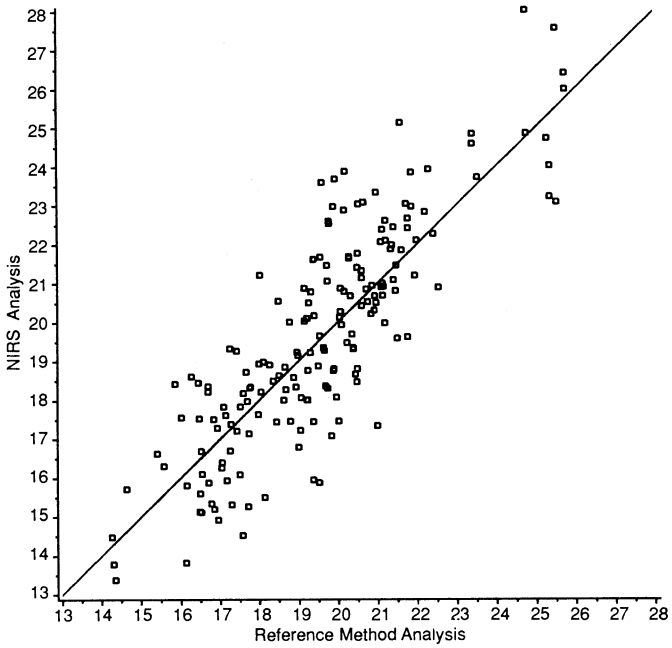


Figure 4.—An unacceptable scatter relationship between NIRS analysis and the reference method.

reference method, or (2) the validation samples are not represented in the calibration population, or (3) the primary reference method data were inaccurate. If less than 10 percent of the validation samples are *H* outliers from the PRE program, one may eliminate the second reason as the probable cause of the high SEP(C).

Print the residual deviations between the predicted values and the primary reference values. If the residuals exhibit a skewed distribution with a few samples having very large residuals, the reference method results for these samples should be suspected. Reanalyze these samples with the reference method. If reference method errors were at fault, use the revised values to determine the SEP. If the original reference analysis was correct, the wavelengths in the equation cannot reliably estimate reference method analyses (that is, the first reason prevails). In this case, recalibration is necessary to include these samples in the calibration population set. If the residuals are large and randomly distributed, they also indicate that recalibration is necessary to include the validation samples in the calibration set. This will necessitate the development of a new sample set for validation of the recalibrated equation. If recalibration is unsuccessful, the precision of the reference method for the validation samples that were added to the calibration set should be suspected. Reanalyze these samples in duplicate by the reference method and compare the results to the original analysis. If the reference method errors were at fault, use the revised values. The need for recalibration may also be evaluated by calculating a bias and SEP control limit as described in supplement 2.

References

Abrams, S. M., J. S. Shenk, M. O. Westerhaus, and F. E. Barton II. 1987. Determination of forage quality by near-infrared reflectance spectroscopy: Efficacy of broad-based calibration equations. J. Dairy Sci. 70:806–813.

Marten, G. C., J. L. Halgerson, and J. H. Cherney. 1983. Quality prediction of small grain forages by near infrared reflectance spectroscopy. Crop Sci. 23:94–96.

Martin, N. P., G. C. Marten, D. A. Schriever, and others. 1985. Near infrared reflectance spectroscopy analysis of quality of farm-grown alfalfa. *In* Proc. 15th Int. Grassl. Cong., Kyoto, Japan, p. 965–966. Sci. Council of Japan/Japanese Soc. Grassl. Sci., Nishi-Nasuno, Tochigi-ken, Japan.

Norris, K. H., R. F. Barnes, J. E. Moore, and J. S. Shenk.
1976 Predicting forage quality by infrared reflectance spectroscopy. J. Anim. Sci. 43:889–897.

Parnell, A., and J. White.
1983 The use of near infrared reflectance analysis in predicting the digestibility and the protein and water soluble carbohydrate contents of perennial ryegrass. J. Natn. Agric. Bot. 16:221–227.

Shenk, J. S., and R. F. Barnes.
1977 Current status of infrared reflectance. *In* Proc. 34th South. Pasture and Forage Crop Improvement Conf., p. 57–62.

Shenk, J. S., M. O. Westerhaus, and M. R. Hoover.
1978 Infrared reflectance analysis of forages. *In* Proc. Int. Grain and Forage Harvesting Conf., p. 242–244, 252. Am. Soc. Agric. Eng., St. Joseph, MI.

Shenk, J. S., M. O. Westerhaus, and M. R. Hoover.
1979 Analysis of forages by infrared reflectance. J. Dairy Sci. 62:807–812.

Windham, W. R., J. A. Robertson, and R. G. Leffler.
1987 A comparison among methods of moisture determination of forages for NIRS calibration and validation. Crop Sci. 27:777–783.

2. Protocol for NIRS Calibration: Monitoring Analysis Results and Recalibration

J. S. Shenk, M. O. Westerhaus, and S. M. Abrams

Monitoring Equation Performance

A system for monitoring the accuracy of an equation for a new group of samples is described. Utilization of this monitoring system on a routine basis will reduce the danger of inappropriate application of calibration equations and aid in laboratory quality assurance.

The monitoring tests and formulas given earlier in this handbook were revised and rewritten for this supplement to be more consistent with standard statistical nomenclature.

After an equation has been selected and validated, continual monitoring is necessary to ensure accurate results. Near infrared reflectance spectroscopy (NIRS) analysis can disagree with reference method values when (1) the instrument is not working properly, (2) the sample population changes and is no longer represented by the calibration set, or (3) some aspect of the reference method changes, usually unknown to the analyst. None of these events must take place if NIRS analysis is to be accurate. The monitoring procedure as presented cannot distinguish between the last two sources of error. If instrument performance is questionable, see "Instrument Operation," p. 24. The major cause of inaccurate results is the application of equations to populations not adequately represented in the group of samples used for calibration. The risks involved have been discussed by Abrams et al. (*1987*), Benson (*1986*), and Minson et al. (*1983*). The unknown samples may be from a different population than the calibration samples because of a difference in processing techniques, chemical composition outside the range of the calibration set, or application of unusual chemical or physical treatments to the samples. During routine NIRS analysis these changes are often not obvious to the technician. Large standardized *H* statistics of individual samples often indicate poor NIRS accuracy (see "Validation," p. 40).

This monitoring system consists of two tests that determine the existence of (1) a significant bias and (2) a significant increase in unexplained error (SEP(C)). If either is occurring, steps should be taken to (1) add samples to the existing calibration set and recalibrate or (2) develop a new calibration for this population based on all new samples.

One sample of every 20 should be routinely set aside for reference method analysis until 9 samples have been accumulated. If predictions of a group of samples appear unreasonable or the standardized *H* values are larger than 3.0, a random set of nine samples should be immediately selected from the group of samples exhibiting the problem. These samples should be analyzed by the reference method and the values compared with the NIRS values by the methods described here. A continuous time chart of the observed bias and unexplained error (SEP(C)) should be maintained to facilitate detection of trends.

A practical monitoring procedure to determine with 90 percent confidence whether a calibration equation is appropriate for a new population of samples is presented here. Statistical tests are constructed for bias as estimated by observed bias, and unexplained error as estimated by SEP(C). We have found that a bias greater than the standard error of calibration (SEC) and an unexplained error greater than 2 times the SEC are unacceptable with calibration equations based on more than 100 samples. For those interested in establishing alternative criteria, see a more complete description in "Statistical Basis for the Monitoring System" (p.105).

Monitoring Procedure

Step 1. Choose nine (*N*) samples from the population to be evaluated. When a prior equation is to be used on a new closed population, nine samples should be selected randomly from the new population. In open populations, such as those encountered routinely in the NIRS laboratory, every 20th sample should be selected until nine samples are accumulated.

Step 2. Obtain analyses of interest by both the reference method and NIRS. Analyze samples by the same reference method used to develop the calibration equation. It is preferred that the same technician perform the analysis. Any bias or increase in unexplained error between NIRS and the reference methods can be due to either inaccuracies in NIRS or changes in reference method procedures.

Step 3. Calculate the bias confidence limits. If the sample size of nine and a probability of Type I error of 0.10 are acceptable, the bias confidence limits are calculated as ± 0.55(SEC).

Step 4. Calculate the unexplained error confidence limit. If a sample size of nine and a probability of Type I error of 0.10 are acceptable, the unexplained error confidence limit is calculated as 1.29(SEC).

Step 5. Determine the observed bias of the test set. Calculate the difference for each of the *N* samples by subtracting the NIRS value from the reference method value. Obtain the average difference by dividing the sum of the differences by *N*. Typically *N* will equal nine.

$$D_i = X_i - Y_i$$

$$\text{Observed bias} = \frac{\Sigma D_i}{N}$$

where D_i = difference, X_i = reference method value for the ith sample, and Y_i = NIRS value for the ith sample.

Step 6. Determine the SEP(C) of the test set. First, square each of the differences calculated in step 5. Subtract N times the square of the observed bias from the sum of the squared differences and divide by $N-1$. The square root of the result is SEP(C).

$$\text{SEP(C)} = \sqrt{\frac{\Sigma (D_i)^2 - N(\text{observed bias})^2}{N-1}}$$

Step 7. Take appropriate action based on the results of steps 5 and 6. If the observed bias and SEP(C) of nine test samples do not remain within the confidence limits, the samples should be added to the calibration set and recalibration should be performed. An exception would be if the observed bias of test samples from a closed homogeneous population exceeded the confidence limits while the unexplained error was within the confidence limits. In this instance, the equation intercept could be adjusted without recalibration.

Practical Example of Monitoring System

A calibration file of 500 hay samples was used to develop an analytical NIRS equation for protein, in vitro dry matter disappearance (IVDMD), and dry matter (Abrams et al. *1987*). These samples were field cured and cyclone ground to pass a 1-mm screen.

Bias confidence limits were set to distinguish between no bias and a bias greater than 1.0 times the SEC with 90 percent confidence ($\beta = 0.10$) when using a two-tailed Type I error probability (α) of 0.10. The unexplained error confidence limit was set to distinguish between no increase in unexplained error and unexplained error that was 2 times SEC or greater with 90 percent confidence ($\beta = 0.10$) and a one-tailed Type I error probability (α) of 0.10.

Three populations of samples not included in the calibration were chosen for this example. Population A was from Pennsylvania and consisted of 100 hay samples of various legume and grass species. These samples were either oven, freeze, or field dried and ground with a Wiley mill to pass a 1-mm screen. Population B consisted of 100 alfalfa hay samples from Michigan that were either oven or field dried and ground with a cyclone mill to pass a 1-mm screen. Population C consisted of 89 grass hay samples from Norway that were field dried and cyclone ground. Nine samples from all three populations were analyzed for protein, IVDMD, and dry matter by both NIRS and the reference method.

Table 1 is the decisionmaking table. Statistics from group A analyses were near or over the confidence limits five out of six times. This was probably due to sample processing differences between the calibration and test samples. Samples from population B could be accurately analyzed by the equations. Both observed bias and SEP(C) were within the confidence limits. Group C had a significant observed bias on both protein and IVDMD. This large IVDMD observed bias could be due to the fact that the in vitro reference method analyses for the calibration samples and population C samples were performed at different times.

All samples in the three populations were analyzed for protein, IVDMD, and dry matter to show that the statistics on the randomly selected subsets ($N = 9$) were similar to the statistics of the entire populations (table 2).

Statistical Basis for the Monitoring System

Bias can be tested with a t test of the following null hypothesis: The average difference between NIRS and reference method values is zero. The unexplained error can be tested with an F test of the following null hypothesis: The unexplained error after correcting for bias is less than or equal to the calibration error.

The unexplained error has two components: the error inherent in the calibration equation (estimated by SEC) and the error resulting when the wavelengths and/or regression coefficients are not correct for the new population. Inappropriate wavelengths result in an increase in unexplained error that can be reduced only through calibration. An incorrect equation intercept (b_0) results in a systematic error—all NIRS analyses being too high or too low, commonly called bias. When the other (nonintercept) regression coefficients ($b_1 \ldots b_n$) are not correct, the resulting NIRS analyses are less accurate. If these coefficients are all too large or too small, the NIRS analyses exhibit a systematic "slope" effect, in which the high values are too high and the low values are too low or vice versa.

It is difficult to establish acceptability levels for slope. The variability of slope depends on the NIRS measurements of the samples. This dependence makes it impossible to guarantee a specific standard error of the slope based solely on the number of samples. For that reason, this monitoring system will address (1) bias and (2)

Table 1.—*Observed bias and SEP(C)[a] of 9 samples with their confidence limits*

Analysis	Population	Hbar[b]	Bias test			Unexplained error test		
			Bias[c]	CL[d]	|Bias| < CL	SEP(C)	CL	SEP(C) < CL
Protein	A	2.45	−1.68	±0.50	No	1.90	1.17	No
	B	.71	−.34	± .50	Yes	1.14	1.17	Yes
	C	2.05	.69	± .50	No	.51	1.17	Yes
IVDMD[e]	A	2.58	2.68	±1.61	No	4.05	3.80	No
	B	1.13	.40	±1.61	Yes	1.54	3.80	Yes
	C	2.17	6.04	±1.61	No	3.27	3.80	Yes
DM[f]	A	2.90	−.23	± .27	Yes	1.09	.64	No
	B	.66	.23	± .27	Yes	.40	.64	Yes
	C	1.00	−.15	± .27	Yes	.65	.64	Yes

[a] Standard error of performance corrected for bias.
[b] Average standardized *H* statistic.
[c] Average difference between reference method and NIRS analytical values.
[d] Confidence limit.
[e] In vitro dry matter disappearance.
[f] Dry matter.

Table 2.—*Statistics for all samples in populations A, B, and C and randomly selected subsets of 9 samples*

Analysis	Population	SEC[a]	(R^2)[b]	N	Mean	SD[c]	Hbar[d]	Bias[e]	SEP(C)[f]	R^2
Protein	A	0.910	0.97	105	16.86	3.87	2.25	−1.15	1.37	0.90
				9	16.86	4.01	2.45	−1.68	1.90	.86
	B	.910	.97	100	18.22	1.97	.77	.25	.91	.79
				9	17.85	2.60	.71	−.34	1.14	.82
	C	.910	.97	89	14.01	2.45	1.91	.69	.67	.93
				9	14.94	2.39	2.05	.69	.51	.97
IVDMD[g]	A	2.946	.88	105	66.76	4.86	2.46	1.73	3.87	.38
				9	68.43	4.80	2.58	2.68	4.05	.39
	B	2.946	.88	100	66.57	3.68	1.21	.25	2.09	.72
				9	64.04	3.68	1.13	.40	1.54	.83
	C	2.946	.88	90	71.89	4.22	2.01	5.91	2.76	.60
				9	73.05	3.27	2.17	6.04	3.27	.23
DM[h]	A	.495	.97	105	94.24	1.62	2.57	−.27	.74	.81
				9	94.51	2.10	2.90	−.23	1.09	.74
	B	.495	.97	100	93.04	.81	.73	.28	.52	.65
				9	93.53	.61	.66	.23	.40	.60
	C	.495	.97	90	93.04	.41	1.04	−.02	.36	.26
				9	92.93	.48	1.00	−.15	.46	.08

[a] Standard error of calibration.
[b] Fraction of explained variance.
[c] Standard deviation.
[d] Average standardized *H* statistic.
[e] Average difference between reference method and NIRS analytical values.
[f] Standard error of performance corrected for bias.
[g] In vitro dry matter disappearance.
[h] Dry matter.

unexplained error as estimated by SEP(C), the slope error remaining unaddressed and contained within the unexplained error.

The analyst must select an alternative hypothesis to the null hypothesis. Confidence limits are established to determine whether the test set values are more likely to arise from the null hypothesis or the alternative hypothesis. An alternative hypothesis can best be expressed as a ratio of bias or unexplained error to true calibration error. These ratios must be preselected by the analyst so that the appropriate number of samples can be selected at random from the population of samples. A bias ratio of 0.5 to 1.0 and an unexplained error ratio of 1.5 to 2.0 are recommended.

The analyst must also decide what levels of Type I and Type II errors are acceptable in the tests. A Type I error for bias occurs when the observed bias exceeds the confidence limits, indicating that the true bias is different from zero when in fact it is not. A Type I error for unexplained error occurs when the SEP(C) exceeds the confidence limit, indicating that the unexplained error is larger than the calibration error when in fact it is not. A Type I error results in making unnecessary corrections to the calibration equation.

A Type II error for bias occurs when the observed bias is within the confidence limits, indicating that the true bias is not different from zero when in fact it is. Similarly, a Type II error for unexplained error occurs when the SEP(C) is less than the confidence limit, indicating that it is not larger than the calibration error although it is. A Type II error results in generating NIRS analytical values with unacceptable errors.

After the alternative hypothesis and levels of Type I and II errors have been selected, the number of samples needed to establish confidence limits can be calculated.

Calculation of Number of Samples for Bias Confidence Limits

The number of samples needed to establish confidence limits for bias is found with the following formula:

$$N > [(t_1 + t_2)/B]^2 \quad (1)$$

where:

- N = number of samples to be used for the test
- t_1 = Student's t value associated with a chosen level of significance for Type I error divided by 2, and the degrees of freedom in the calibration sample set
- t_2 = Student's t value associated with a chosen level of significance for a Type II error, and the degrees of freedom in the calibration sample set
- B = alternative hypothesis, expressed as the ratio of bias to true calibration error

Example 1: Determine the number of samples required to establish bias confidence limits for an NIRS calibration equation based on 200 samples to accurately analyze a new population.

1. Select an alternative hypothesis. In this example, the analyst wants to distinguish between no bias and a bias larger than the true calibration error such that the bias ratio is 1.0 (B = 1.0). The SEC in this example is 2.4.

2. Establish acceptable levels of Type I (α) and Type II (β) errors. In this example, $\alpha = 0.10$ and $\beta = 0.10$. The α value is divided by 2 because it is a two-tailed test.

The α or β probability points of Student's t distribution for infinite degrees of freedom are as follows:

α *or* β	t
0.25	0.67
.100	1.282
.050	1.645
.025	1.960
.010	2.326
.005	2.576

3. From these data substitute t values in formula (1):

$t\,(\alpha/2) = t_1 = 1.645$
$t\,(\beta) = t_2 = 1.282$
$N > [(1.645 + 1.282)/1.0]^2$
$N > [2.927]^2$
$N > 8.567$; therefore, 9 samples are required.

Example 2: Using the information given in example 1, determine the number of samples required for the alternative hypothesis in which the bias ratio is 0.5 (B = 0.5).

$N > [(1.645 + 1.282)/0.5]^2$
$N > [(2.927)/0.5]^2$
$N > [5.854]^2$
$N > 34.269$; therefore, 35 samples are required.

Calculation of Bias Confidence Limits

Determine the confidence limits for accepting or rejecting equation performance on the small set of samples chosen from the new population.

$$\text{Bias confidence limits} = \pm (t_1 \cdot \text{SEC})/\sqrt{N} \qquad (2)$$

where:

t_1 = appropriate Student's t value from the preceding tabular data for a two-tailed test with degrees of freedom associated with SEC and the selected probability of a Type I error

N = number of samples (formula (1)) chosen from the new population to be analyzed by the reference method

For example 1 with nine samples and a probability of Type I error of 0.10/2, the result is—

$$\text{Bias confidence limits} = \pm (1.645 \cdot 2.4)/\sqrt{9}$$

$$= \pm (3.948)/3 = \pm 1.32$$

If the observed bias of the test set is between -1.32 and $+1.32$, it is within the bias confidence limits.

Calculation of the Number of Samples for Unexplained Error Test

The number of samples needed to establish the confidence limit for unexplained error is found with the following formulas:

$$\frac{F_{(\alpha:n,m)}}{F_{(1-\beta:n,m)}} < A^2 \qquad (3)$$

$$N = n + 1$$

where:

α = probability of making a Type I error
β = probability of making a Type II error
n = numerator degrees of freedom associated with SEP(C) of test set
m = denominator degrees of freedom associated with SEC
A = alternative hypothesis, expressed as ratio of unexplained error of the test population to the unexplained error of the calibration population. Note that SEP(C) and SEC are only estimates of these parameters.
N = minimum number of samples needed to establish the confidence limit.

Note that the degrees of freedom n of SEP(C) is $N-1$, so that $N = n + 1$. Most readily available F tables do not give the lower probability points. The standard tables of upper probability points may be used with this substitution:

$$F_{(\beta:m,n)} = \frac{1}{F_{(1-\beta:n,m)}}$$

giving this formula:

$$F_{(\alpha:n,m)} \cdot F_{(\beta:m,n)} < A^2 \qquad (4)$$

Table 3 provides selected probability levels of the F distribution.

Example 3: Determine the number of samples required to adequately test the unexplained error of a calibration equation based on 200 samples to analyze a new group of 100 samples.

1. Select an alternative hypothesis. In this example, the analyst wants to distinguish between unexplained error equal to true calibration error and unexplained error twice as large as true calibration error (A = 2.0). The SEC error in this example is 2.4.

2. Establish Type I and Type II error limits. In this example $\alpha = 0.10$ and $\beta = 0.10$.

3. Obtain the minimum number of samples N to establish the confidence limit by first finding the smallest degrees of freedom n in formula (3), so that the product of the two F values is less than A^2, and then adding one to get the minimum number of samples N. Use table 4 to obtain N when $\alpha = 0.10$, $\beta = 0.10$, and degrees of freedom associated with SEC is approximated by infinity.

4. Find the first number in the A' column less than the A of the alternative hypothesis ($A' < 2$). The number of degrees of freedom in that row is the degrees of freedom needed in the SEP(C) to construct the confidence limit. In this example, eight degrees of

Table 3.—α *and* β *probability points of central* F *distribution with* n *degrees of freedom of SEP(C) and* m *degrees of freedom of calibration error (SEC)*

α or β	5	6	7	8	9	10	12	15	20	30	60	120
						($n,m = \infty$)						
0.10	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.34	1.24	1.17
.05	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.32	1.22
.025	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.57	1.39	1.27
						($m = \infty, n$)						
.10	3.11	2.72	2.47	2.29	2.16	2.06	1.90	1.76	1.61	1.46	1.29	1.19
.05	4.37	3.67	3.23	2.93	2.71	2.54	2.30	2.07	1.84	1.62	1.39	1.25
.025	6.02	4.85	4.14	3.67	3.33	3.08	2.72	2.40	2.09	1.79	1.48	1.31

Table 4.—*Appropriate number of samples to use to establish confidence limits*

N^a	n^b	$F_{(0.10:n,\infty)}{}^c$		$F_{(0.10:\infty,n)}{}^d$		$(A')^2$	A'^e
7	6	1.77	×	2.72	=	4.814	2.194
8	7	1.72	×	2.47	=	4.248	2.061
9	8	1.67	×	2.29	=	3.824	1.955
10	9	1.63	×	2.16	=	3.521	1.876
11	10	1.60	×	2.06	=	3.296	1.815
13	12	1.55	×	1.90	=	2.945	1.716
16	15	1.49	×	1.76	=	2.622	1.619
21	20	1.42	×	1.61	=	2.286	1.512
22	21	1.41	×	1.59	=	2.237	1.496
26	25	1.38	×	1.52	=	2.087	1.445
31	30	1.34	×	1.46	=	1.956	1.399

[a] Minimum number of samples needed to establish the confidence limit.

[b] Degrees of freedom associated with SEP(C).

[c] *F* value from *F* distribution table when α = 0.10, β = 0.10 and the denominator degees of freedom associated with SEC is approximated by infinity.

[d] *F* value from *F* distribution table when α = 0.10, β = 0.10 and the numerator degees of freedom associated with SEC is approximated by infinity.

[e] Ratio of unexplained error to error of calibration.

freedom (n = 8), or nine samples (N = 9), are required to establish a confidence limit with A = 2.0. But another analyst might prefer, for example, to distinguish between unexplained error equal to true calibration error and unexplained error 1.5 times as large as true calibration error (A = 1.5). The value for n would then be 21 and that for N would be 22.

Calculation of Unexplained Error Confidence Limits

Confidence limits for unexplained error are calculated with the following formula:

$$\text{Unexplained error confidence limit} = \text{SEC} \cdot \sqrt{[F_{(\alpha:n,m)}]} \quad (5)$$

Example 4: Determine the unexplained error confidence limit for the previous example in which N = 9, A = 2.0, and SEC = 2.4.

$$\text{Unexplained error confidence limit} = 2.4 \cdot \sqrt{(1.67)} = 3.10$$

If the SEP(C) of the test set is larger than 3.10, the equation will probably not perform satisfactorily on the set of 100 samples.

Recalibration

Recalibration is the process of adding new samples to the existing calibration set and deriving a new calibration equation. This is necessary whenever (1) the NIRS monitoring results fall outside the confidence limits, (2) many samples exhibit large H values, or (3) the instrument environment changes.

Additional samples for recalibration should be selected on the basis of spectral characteristics. Samples for recalibration should be chosen with the same program used to select samples for calibration (see "Calibration Sample Selection," p. 96). New samples are compared to the existing samples and classified as (1) similar to one or more existing samples (do not add), (2) not similar to any existing sample (add), and (3) different (unconnected and possibly from another population) from all existing samples (do not add).

The added samples should be analyzed by the primary reference method and the augmented set of calibration samples resubmitted to the entire calibration procedure. After the "best" equation for each constituent has been selected, the performance of this new equation should continue to be monitored. It is often necessary to repeat the calibration, monitoring, and recalibration cycle a number of times to obtain a good robust calibration.

References

Abrams, S.M., J.S. Shenk, M.O. Westerhaus, and F.E. Barton II. 1987. Determination of forage qualilty by near infrared reflectance spectroscopy: Efficacy of broad-based calibration equations. J. Dairy Sci. 70:806-813.

Benson, R.D. 1986. Statistical process control techniques using near-infrared reflectance spectroscopy. *In* A.R. Baldwin, ed., Proceedings of World Conference on Emerging Technologies in the Fats and Oils Industry. Cannes, France, Nov. 3-8, 1985. American Oil Chemists Society, Champaign, IL. p. 257-260.

Minson, D.J., K.L. Butler, N. Grummitt, and D.P. Law. 1983. Bias when predicting crude protein, dry matter digestibility and voluntary intake of tropical grasses by near-infrared reflectance. Anim. Feed Sci. and Tech. 9:221-237.